AD-755 797

BAYESIAN DECISION THEORY APPLIED TO THE FINITE STATE MARKOV DECISION PROBLEM

William Ross Osgood

California University

Prepared for:

Air Force Office of Scientific Research
Air Force Space and Missile Systems Organization

September 1972

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

UNIVERSITY OF CALIFORNIA
ENGINEERING · LOS ANGELES

UCLA-ENG-7278
SEPTEMBER 1972

BAYESIAN DECISION THEORY APPLIED TO THE FINITE STATE MARKOV DECISION PROBLEM

W.R. OSGOOD

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of California School of Engineering and Applied Science Los Angeles, California 90024	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

BAYESIAN DECISION THEORY APPLIED TO THE FINITE STATE MARKOV DECISION PROBLEM

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

W. R. Osgood

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1972	100	14

8a. CONTRACT OR GRANT NO. AFOSR 72-2166 b. PROJECT NO. 9769 c. 61102F d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) UCLA-ENG-7278 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR - TR - 73 - 0228

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH OTHER	Air Force Office of Scientific Rsch (NM) 1400 Wilson Blvd Arlington, Virginia 22209

13. ABSTRACT

Ron Howard solved the Markov decision problem with perfect knowledge of all the transition probabilities and rewards. In a practical situation, the transition probabilities may not be known exactly. Therefore, the problem this research attacks is the Markov decision problem with uncertain transition probabilities.

In the case of perfect knowledge, the decision that maximizes the expected reward or gain is chosen. When there is uncertainty in the transition probabilities, the gains become random variables. Therefore, Bayesian decision theory is applied to this problem. A loss function is defined and an a priori density is defined. Bayes' formula and the loss function are used to compute a risk for each decision. The decision that minimizes the risk is chosen.

Conceptually the problem is solved easily. However, transforming a density over the transition probabilities to a density over the gains is a difficult problem. The solution of this problem is the main contribution of this dissertation. Using these results a technique is derived that allows a straightforward means to evaluate the risks for each decision. Examples are presented that illustrate the technique.

The result of this research is a logical method to compute the risks associated with each decision when there is uncertainty over the transition probabilities. The decision maker then selects the decision that minimizes the risk.

DD FORM 1 NOV 65 1473

Ia

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Decision Theory						
Markov Processes						
Bayesian Decision Theory						

ib

Security Classification

UCLA-ENG-7278
September 1972

BAYESIAN DECISION THEORY APPLIED TO THE FINITE STATE MARKOV DECISION PROBLEM

W.R. Osgood

School of Engineering and Applied Science
University of California
Los Angeles

PREFACE

Decision making is an issue constantly before either the developer or user of U.S. Air Force space, missile, tactical, or other systems. Yet since Howard's significant work of over 12 years ago there has been little progress in this area on the important methods pioneered by Howard. Because of the importance of this area to applied Air Force needs the numerous results embodied in this research report were developed and illustrated through numerous examples presented herein.

This research report was prepared under research contracts supported by the U.S. Air Force Office of Scientific Research under AFOSR Grant No. 72-2166, Design of Aerospace Systems, and the U.S. Air Force Space and Missile Systems Organization under Contract No. F04701-72-C-0273, Advanced Space Guidance, and this report constitutes part of the final report under these contracts.

The research described in this report "Bayesian Decision Theory Applied To The Finite State Markov Decision Problem," UCLA-ENG-7278, by William Ross Osgood, was carried out under the direction of C.T. Leondes and E.B. Stear, Co-Principal Investigators in the Schools of Engineering in the University of California at Los Angeles and Santa Barbara, respectively.

ABSTRACT

Ron Howard solved the Markov decision problem with perfect knowledge of all the transition probabilities and rewards. In a practical situation, the transition probabilities may not be known exactly. Therefore, the problem this research attacks is the Markov decision problem with uncertain transition probabilities.

In the case of perfect knowledge, the decision that maximizes the expected reward or gain is chosen. When there is uncertainty in the transition probabilities, the gains become random variables. Therefore, Bayesian decision theory is applied to this problem. A loss function is defined and an a priori density is defined. Bayes' formula and the loss function are used to compute a risk for each decision. The decision that minimizes the risk is chosen.

Conceptually the problem is solved easily. However, transforming a density over the transition probabilities to a density over the gains is a difficult problem. The solution of this problem is the main contribution of this dissertation. Using these results a technique is derived that allows a straightforward means to evaluate the risks for each decision. Examples are presented that illustrate the technique.

The result of this research is a logical method to compute the risks associated with each decision when there is uncertainty over the transition probabilities. The decision maker then selects the decision that minimizes the risk.

CONTENTS

 Preceding page blank

CONTENTS (Continued)

FIGURES

Chapter 1

PROBLEM DEFINITION

The objective of this research is to apply Bayesian decision theory to the finite state Markov decision problem when the transition probabilities are unknown. A Markov decision problem exists when a decision maker has available a set of K decisions. Each decision specifies a particular Markov chain and a set of rewards. The decision maker selects the decision that maximizes his gain or expected reward.

Bayesian decision theory can be used when there is uncertainty over the transition probabilities. An a priori density is specified over the probabilities and Bayes' formula is used to compute an updated posteriori density after observations are recorded. The decision maker selects the decision that minimizes his expected loss or risk. As more observations are recorded the posteriori density concentrates its probability mass over the actual values of the transition probabilities and the risk minimizing decision maximizes the gain.

It is assumed that the reader has knowledge of the theory of Markov chains (see Reference [7]). The following notation is used in this dissertation. There are N states in the Markov chain under consideration. The probability of making a transition from state i to state j is denoted by p_{ij}. The NxN transition matrix is denoted by P. The steady state probability vector is denoted by $\underline{\pi}$.

A knowledge of Bayesian decision theory is also assumed (See Reference [3]). The following notation is used. The states of nature is denoted by Ω . The observation is written as $\underline{x}_n$. The a priori density over Ω is $\xi_0(w)$ where w is an element of Ω. The posteriori density is denoted by $\xi(w|\underline{x}_n)$ and is computed from Bayes' formula,

$$\xi(w|\underline{x}_n) = \frac{l(\underline{x}_n|w)\ \xi_0(w)}{\int_{\Omega} l(\underline{x}_n|w)\ \xi_0(w)\ dw}$$

where $l(\underline{x}_n|w)$ is the likelihood function. For each element w of Ω and each decision k, a loss $L(k|w)$ is incurred. The risk $\rho(k)$ is the expected loss.

$$\rho(k) = \int_{\Omega} L(k|w)\ \ \xi(w|\underline{x}_n)\ dw$$

Bayes decision k* minimizes the risk,

$$\rho(k^*) = \min_i \{\rho(i)\}$$

1.1 The Markov Decision Problem

Once a Markov chain is defined, a reward structure can be placed over the states. Suppose that payoff r_i is received when state i is occupied. The N-vector $\underline{r} = (r_1, \ldots, r_N)$ is called the reward vector associated with the N-state Markov chain. In steady state, the expected payoff or gain is denoted by Δ where

$$\Delta = \sum_{i=1}^{N} r_i \pi_i$$
$$= \langle \underline{r}, \underline{\pi} \rangle$$

Now, suppose that there are K decisions available to a decision maker. Each decision i, $i = 1, \ldots, K$, specifies a unique N-state Markov chain with transition matrix P_i. The corresponding reward vector is denoted by $\underline{r}^i$. The gain under decision i is denoted by Δ_i where

$$\Delta_i = \langle \underline{\pi}^i, \underline{r}^i \rangle$$

and

$$\underline{\pi}^i = \underline{\pi}^i P_i$$

The decision maker selects the decision that maximizes his expected payoff or gain. In otherwords, he will select decision k* such that

$$\Delta_{k*} = \max_{1 \leq i \leq K} \{\Delta_i\}$$

1.2 The Markov Decision Problem with Uncertainty

The Markov decision problem defined above assumes that all transition probabilities and rewards are known with certainty. However, there may be a case where there is uncertainty in some or all of the transition probabilities and/or rewards. The case of perfect information was developed by Ron Howard [7] in 1960. After Howard completed his work others at MIT continued to investigate this problem with uncertainties. The goal of these works should have been to specify the "best" decision against some criterion. However,

their work, summarized by Martin [9] in 1967, did not include a means to specify a decision.

This research applies Bayesian decision theory to the Markov decision problem so that a decisior can be specified under uncertainty. The transition probabilities are taken as uncertain, but the rewards are assumed known.

Martin showed that if the states of nature are the set of all possible transition matrices, and the matrix beta density is used as the a priori density, then Bayes formula transforms observations of state transitions into a posteriori density that is also matrix beta. Therefore, the states of nature is defined as

$$\Omega = \left\{ \Lambda_1^N, \ \Lambda_2^N, \dots \Lambda_K^N \right\}$$

where Λ_i^N is the set of all possible NxN transition matrices under decision i. An element of Ω is denoted by w where

$$w = (P_1, \ P_2, \dots, \ P_K)$$

and P_i is the transition matrix under decision i.

The end product in applying Bayesian decision theory is a risk associated with each decision. The risk is defined as the expected loss. If $L(i|w)$ associates a loss to each decision i when $w \in \Omega$ is the state of nature, then the risk $\rho(i)$ becomes

$$\begin{aligned} \rho(i) &= E \ L(i|w) \\ &= \int_\Omega L(i|w) \, \xi(w|\underline{x}_n) \, dw \end{aligned}$$

where $\xi(w|\underline{x}_n)$ is the posteriori density over Ω. The decision maker's objective is to maximize his gain, therefore the loss will be defined in terms of the gain. Take element $w \in \Omega$ where

$$w = (P_1, \ldots, P_K)$$

To each P_i the steady state probability vector $\underline{\pi}^i$ can be computed. Then, for each decision i the gain Δ_i is given by

$$\Delta_i = < \underline{\pi}^i, \underline{r}^i >$$

Suppose decision i is selected resulting in a payoff of Δ_i units per transition. But, if there is a Δ_j $j = 1, \ldots, K$ such that $\Delta_j > \Delta_i$ then the decision maker suffers a loss of at least $\Delta_j - \Delta_i$ per transition because he selected decision i instead of decision j. The loss function $L(i|w)$ is defined as the maximum loss and is given by

$$L(i|w) = \max_{1 \le j \le K} \{\Delta_j - \Delta_i\}$$

Thus, for each $w \in \Omega$, $L(i|w)$ is the maximum loss per transition when decision i is chosen.

In order to calculate the risk, the expression $\max\limits_{1 \le j \le K} \{\Delta_j - \Delta_i\}$ must be written in terms of $w = (P_1, \ldots, P_K)$. Gain Δ_i is written as

$$\Delta_i = < \underline{\pi}^i, \underline{r}^i >$$

The problem with this expression is that $\underline{\pi}^i$ must be specified as an explicit function of P_i. Since the steady state probability vector $\underline{\pi}^i$ is uniquely related to its transition matrix P_i by $\underline{\pi}^i = \underline{\pi}^i P_i$

the existence of a function $\underline{\pi}^i = \underline{\sigma}(P_i)$ has some intuitive appeal.

Now, gain Δ_i is written as

$$\Delta_i = \langle \underline{\sigma}(P_i),\ \underline{r}^i \rangle$$

This expression is substituted to get the desired expression

$$L(i|w) = \max_{1 \leq j \leq K} \left\{ \langle \underline{\sigma}(P_j),\ \underline{r}^j \rangle - \langle \sigma(P_i),\ \underline{r}^i \rangle \right\} \tag{1.1}$$

The risk function becomes

$$\rho(i) = \int_{\Omega} \max_{1 \leq j \leq K} \left\{ \langle \underline{\sigma}(P_j), \underline{r}^j \rangle - \langle \underline{\sigma}(P_i), \underline{r}^i \rangle \right\} \xi(P_1, \ldots, P_K | \underline{x}_n) dw$$

$$i = 1, \ldots, K \quad (1.2)$$

The decision maker continues to make observations $\underline{x}_n$ until he is satisfied that the posteriori density has concentrated a sufficient amount of probability mass over the actual collection of transition matrices. Then the risk is computed. The decision that minimizes the risk is chosen.

The problem is to evaluate Equation 1.2. In picking the risk minimizing decision, the absolute value of the risk is not important. That is, the decision the results in the smallest risk, whatever its value, is chosen. Suppose that risk $\rho(i)$ is used as a reference. The minimum risk $\rho(k^*)$ will satisfy,

$$\rho(k^*) - \rho(i) \leq \rho(k) - \rho(i) \qquad k=1, \ldots, K \tag{1.3}$$

An alternate way of expressing Equation 1.3 is to say that k^* satisfies the expression,

$$\rho(k^*) - \rho(i) = \min_k \left\{ \rho(k) - \rho(i) \right\}$$

Now, the expression $\rho(i) - \rho(k)$ is evaluated by substituting Equation 1.2,

$$\rho(i) - \rho(k) = \int_\Omega \max_j \left\{ <\underline{\sigma}(P_j), \underline{r}^j> - <\underline{\sigma}(P_i), \underline{r}^i> \right\} \xi(w|\underline{x}_n)\, dw$$

$$- \int_\Omega \max_j \left\{ <\underline{\sigma}(P_j), \underline{r}^j> - <\underline{\sigma}(P_k), \underline{r}^k> \right\}$$

$$\xi(w|\underline{x}_n)\, dw$$

$$= \int_\Omega (<\underline{\sigma}(P_k), \underline{r}^k> - <\underline{\sigma}(P_i), \underline{r}^i>)\, \xi(w|\underline{x}_n)\, dw$$

$$= \int_\Omega <\underline{\sigma}(P_k), \underline{r}^k> \xi(w|\underline{x}_n)\, dw - \int_\Omega <\underline{\sigma}(P_i), \underline{r}^i> \xi(w|\underline{x}_n)\, dw$$

$$= E(\Delta_k|\underline{x}_n) - E(\Delta_i|\underline{x}_n) \qquad (1.4)$$

Thus, expression $\rho(i) - \rho(k^*)$ is given by

$$\rho(i) - \rho(k^*) = \min_k \left\{ E(\Delta_k|\underline{x}_n) - E(\Delta_i|\underline{x}_n) \right\} \qquad (1.5)$$

Define function $\eta^i(k)$ by

$$\eta^i(k) = E(\Delta_i|\underline{x}_n) - E(\Delta_k|\underline{x}_n) \qquad (1.6)$$

The decision k^* that minimizes the risk minimizes $\eta^i(k)$,

$$\eta^i(k^*) = \min_k \left\{ \eta^i(k) \right\}$$

Therefore, the problem of evaluating the risk for each decision is transformed to a problem of evaluating function $\eta^i(k)$ for each decision

$k=1, \ldots, K$.

The expected value of the gain $E(\Delta_k | \underline{x}_n)$ $k=1, \ldots, K$ cannot be computed is a straightforward manner. The procedure required is developed by first looking at a general case in Chapter 2 where there is one Markov chain (one gain) and the uncertainty over the transition matrix is a general function $h_p(P)$. The expected value of the gain is $E(\Delta)$. In Chapter 3 the density over the transition matrix is defined as matrix beta and the observation $\underline{x}_n$ is used to compute the posteriori density $\xi(P | \underline{x}_n)$. The expected value of the gain conditioned on the observation is $E(\Delta | \underline{x}_n)$. In Chapter 4 decisions are introduced. The results of Chapter 3 are used to compute $E(\Delta_k | \underline{x}_n)$ for each decision $k=1, \ldots, K$. The risk minimizing decision k^* is selected by using Equation 1.6. An example is presented that illustrates the procedure developed.

Two secondary issues are discussed in Chapters 5 and 6. In Chapter 5 the possibility of specifying the uncertainty over the steady state probabilities rather than the transition probabilities is explored. In Chapter 6 the theory used to evaluate the expected value of the gains is used to investigate the convergence properties of two state Markov chains.

The results of this research are summarized in Chapter 7 and several topics for future research are outlined.

Chapter 2

COMPUTING THE EXPECTED VALUE OF THE GAIN

The objective of this Chapter is to compute the expected value of the gain $E(\Delta)$ for the general case where there is one Markov chain and the uncertainty over the transition matrix is given by density $h_{\underline{P}}(P)$. The gain is

$$\Delta = \langle \underline{\pi}, \underline{r} \rangle$$

$$= \langle \underline{\sigma}(P), \underline{r} \rangle$$

where the existence of function $\underline{\sigma}(\cdot)$ is hypothesized. The expected value of the gain could be expressed as

$$E(\Delta) = \int_{\Lambda^N} \langle \underline{\sigma}(P), \underline{r} \rangle \, h_{\underline{P}}(P) \, dP \qquad (2.1)$$

where Λ^N is the set of all possible NxN transition matrices. Even though Equation 2.1 looks simple enough there is a serious complication. That is, set Λ^N is not a closed and convex subset of Eucledian space. This complicates the integration operation. What is required is some transformation that allows integration in Eucledian space.

The transition matrix P has N rows, each of which are probability vectors. The i^{th} row is called the i^{th} transition vector and is denoted by $\underline{p}_i$. Suppose that $\underline{p}_i$ is known with certainty. Clearly, this knowledge conveys no information about $\underline{p}_1, \ldots, \underline{p}_{i-1}, \underline{p}_{i+1}, \ldots \underline{p}_N$. In other words, each row of P should be probabilistically independent. Therefore, density $h_{\underline{P}}(P)$ can be written as,

$$h_{\underline{P}}(P) = h_1(\underline{p}_1)\ldots h_N(\underline{p}_N)$$

where $h_i(\underline{p}_i)$ is a density function over the set Π_N, the set of all N dimensional probability vectors. Equation 2.1 becomes

$$E(\Delta) = \int_{\Pi_N x \ldots x \Pi_N} \langle \underline{\sigma}(P), \underline{r} \rangle h_1(\underline{p}_1)\ldots h_N(\underline{p}_N)\, d\underline{p}_1 \ldots d\underline{p}_N \quad (2.2)$$

To see how a probability density $h_i(\underline{p}_i)$ over Π_N can be transformed to a density over a closed convex set, a two dimensional case is examined.

Consider a two dimensional probability vector $\underline{p} = (p_1, p_2)$. The density over $\underline{p}$ is denoted by $h_{\underline{p}}(p_1, p_2)$. Since p_1 and p_2 are constrained to satisfy the conditions,

i) $p_1, p_2 \geq 0$

ii) $p_1 + p_2 = 1$

the probability mass must lie over the line $p_1 + p_2 = 1$ in the positive quatrant as drawn in Figure 2.1. A simpler way to characterize the density would be to define it as a function of a single variable s where $s = 0$ corresponds to $\underline{p} = (1, 0)$ and $s = \sqrt{2}$ corresponds to $\underline{p} = (0, 1)$, for example. In this case the density over $\underline{p}$ is denoted by $f_{\underline{p}}(s)$ and is drawn in Figure 2.2. Notice that density $h_{\underline{p}}(p_1, p_2)$ is transformed to density $f_{\underline{p}}(s)$ where s belongs to the closed convex set

$$\{s \in E^1 \mid 0 \leq s \leq \sqrt{2}\}$$

Vector $\underline{p}$ and scalar s are related through the equation

$$\underline{p} = (1, 0) + \frac{s}{\sqrt{2}}(-1, 1)$$

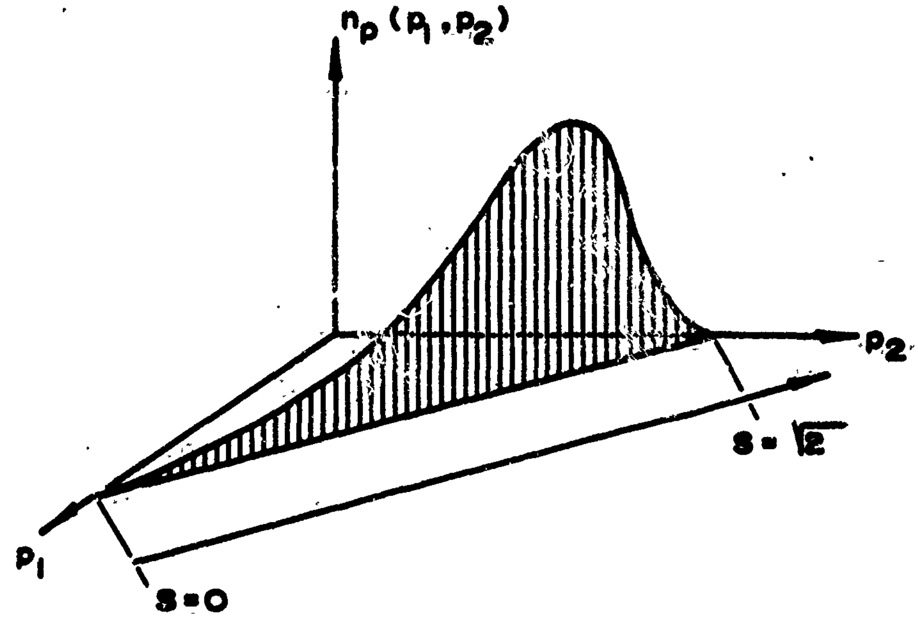


Figure 2.1 Density $f_p(p_1, p_2)$

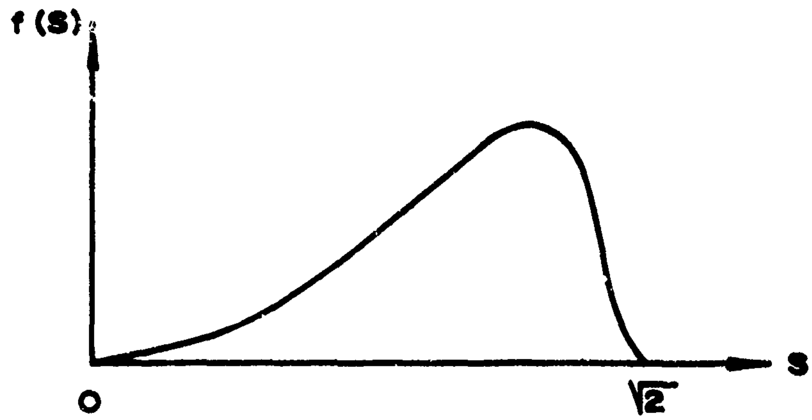


Figure 2.2 Density $h_p(s)$

Now consider a two-dimensional transition matrix P. The density over the first row is denoted by $f_1(s_1)$, and the density over the second row is denoted by $f_2(s_2)$. Since the rows of matrix P are probabilistically independent, the density over matrix P, denoted by f_P, is given by

$$f_P(\underline{t}) = f_1(s_1)\, f_2(s_2)$$

where

$$\underline{t} = (s_1, s_2)$$

A density $f_P(\underline{t})$ is drawn in Figure 2.3. Notice again that the new density $f_P(\underline{t})$ is defined over a closed convex subset of E^2.

This notion of expressing a two dimenstional probability vector as a scalar is extended to the N state case in this Chapter. Density $h_i(\underline{p}_i)$ is easily transformed to density $f_i(\underline{s}_i)$ where $\underline{s}_i$ is a member of set T_N which is a subset of E^{N-1}. Density $h_P(P)$ is transformed to density $f_P(\underline{t})$ where

$$f_P(\underline{t}) = f_1(\underline{s}_1) \cdots f_N(\underline{s}_N)$$

$$\underline{t} = (\underline{s}_1, \ldots, \underline{s}_N)$$

and $\underline{t}$ is a member of set $T_N^N = T_N x \ldots x T_N$ (N times). It will be shown that Equation 2.2 can be written as

$$E(\Delta) = \int_{T_N^N} \langle \underline{w}(\underline{t})U, \underline{r} \rangle f_P(\underline{t})\, d\underline{t} + \langle \underline{\pi}_o, \underline{r} \rangle \qquad (2.3)$$

This expression can be evaluated either analytically for special Markov chains or numerically on the computer.

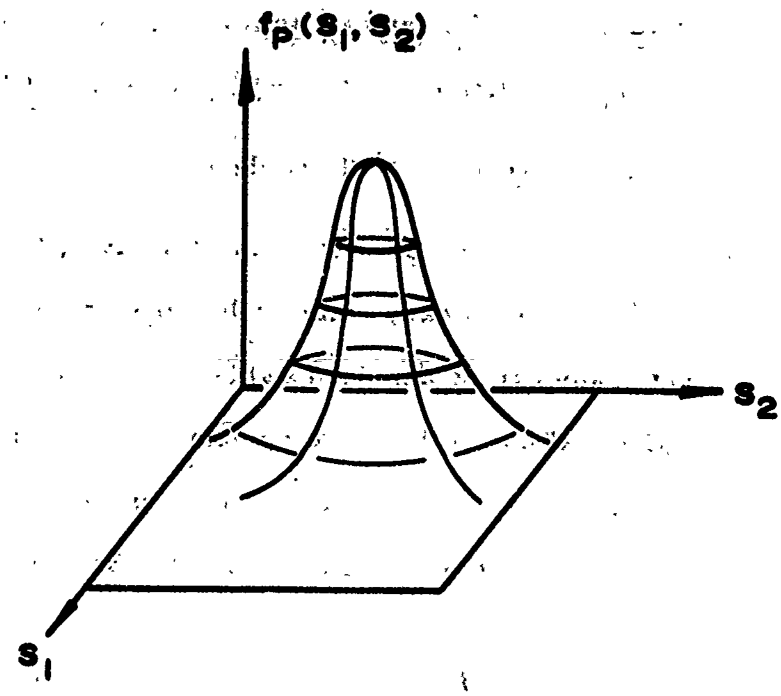


Figure 2.3 Density $f_p(\underline{t})$

In order to evaluate Equation 2.3, function $\underline{w}(\underline{t})$ must be derived, density $f_P(\underline{t})$ has to be transformed from $h_P(P)$, and matrix U and vector $\underline{\pi}_o$ have to be defined. The steps taken to evaluation Equation 2.3 are,

1) Sets T_N and T_N^N are defined.

2) The transformations that take probability vector $\underline{p}$ in Π_N to vector $\underline{s}$ in T_N, and transition matrix P in Λ^N to vector $\underline{t}$ in T_N^N are defined. Matrix U and vector $\underline{\pi}_o$ are defined.

3) Functions $\underline{w}(\underline{t})$ and $\underline{\sigma}(P)$ are derived.

4) Given a transition matrix P and density $h_P(P)$, the density $h_\pi(\underline{\pi})$ over the steady-state probability vector is sought. The transformed density $f_\pi(\underline{s})$ is computed from density $f_P(P)$. Although this development is not used in evaluating Equation 2.3 it is included because of its fundamental importance. Equation 2.3 could also be written as

$$E(\Delta) = \int_{T_N} < \underline{s}U, \underline{r} > f_\pi(\underline{s})\, d\underline{s} + < \underline{\pi}_o, \underline{r} >$$

5) Equation 2.3 is evaluated.

2.1 Sets T_N and T_N^N

All vectors in Π_N have the following property. Given any two probability vectors $\underline{p}_1$ and $\underline{p}_2$, the difference vector $\underline{p}_1 - \underline{p}_2$ lies in the hyperplane H defined by

$$H = \left\{ \underline{x} \in E^N \mid < \underline{x}, \underline{e} > = 0 \right\}$$

Conceptually, probability vectors "touch" hyperplane H as shown in the three dimensional example drawn in Figure 2.4. Since every probability vector can be uniquely represented by a point in H, the set of all these points is a closed convex subset of E^{N-1}. H is an N-1 dimensional hyperplane, therefore a basis χ can be constructed in H where

$$\chi = \left\{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_{N-1}\right\}$$

Then any vector $\underline{z} \in H$ is given by

$$\underline{z} = \sum_{i=1}^{N-1} s_i \, \underline{u}_i$$

Vector $\underline{z}$ with respect to basis χ is given by the N-1 dimensional vector $\underline{s}$ where

$$\underline{s} = (s_1, s_2, \ldots, s_{N-1})$$

Now, set χ together with any vector $\underline{\pi}_o$ in Π_N is clearly a basis of E^N.

Since $\{\chi, \underline{\pi}_o\}$ is a basis in E^N, every vector $\underline{p} \in \Pi_N$ can be written as

$$\underline{p} = \sum_{i=1}^{N-1} s_i \, \underline{u}_i + s_N \, \underline{\pi}_o$$

Vector $\underline{p} - s_N \underline{\pi}_o$ lies in H so

$$\begin{aligned} 0 &= \langle \underline{p} - s_N \, \underline{\pi}_o, \, \underline{e} \rangle \\ &= \langle \underline{p}, \underline{e} \rangle - s_N \langle \underline{\pi}_o, \underline{e} \rangle \\ &= 1 - s_N \end{aligned}$$

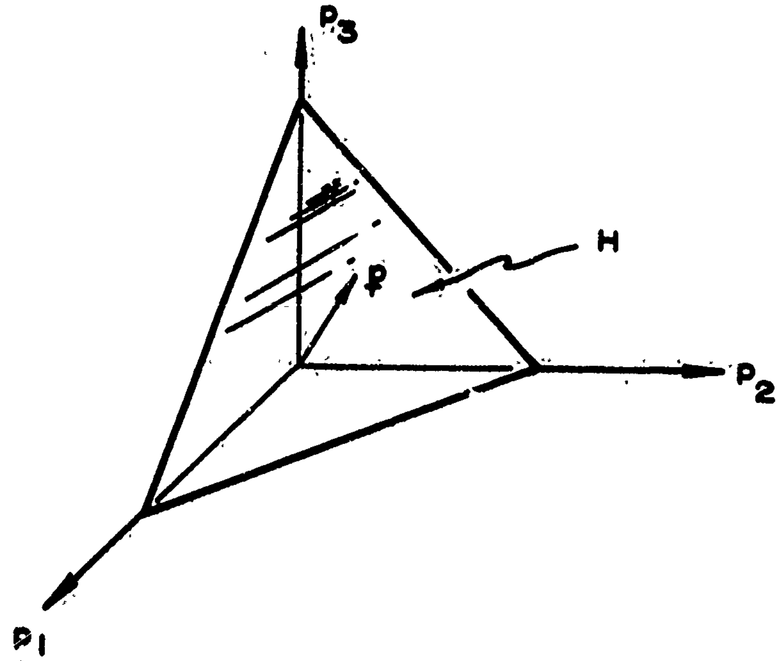


Figure 2.4 Set H in E^3

Therefore, $s_N = 1$ and vector $\underline{p}$ with respect to basis $\{\chi, \pi_o\}$ is given by

$$(s_1, s_2, \ldots, s_{N-1}, 1)$$

Since the N^{th} component is always unity, a probability vector $\underline{p}$ will have a unique representation in basis χ given by the N-1 dimensional vector $\underline{s}$ where

$$\underline{s} = (s_1, \ldots, s_{N-1})$$

The set of all probability vectors is given by Π_N. To each element $\underline{p}$ of Π_N there is a unique $\underline{s}$ in set E^{N-1}. The collection of $\underline{s}$ representing all $\underline{p} \in \Pi_N$ is denoted by T_N where

$$T_N = \left\{\underline{s} \in E^{N-1} \mid \underline{s}U = \underline{p} - \underline{\pi}_o, \underline{p}, \underline{\pi}_o \in \Pi_N\right\} \qquad (2.4)$$

An NxN transition matrix P can be represented by the N^N vector in $E^N x \ldots x E^N$,

$$(\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_N)$$

where $\underline{p}_i$ is the i^{th} row of matrix P. Since each $\underline{p}_i$ is a vector in Π_N, it has a unique representation $\underline{s}_i$ in T_N. Therefore, matrix P has a unique representation in set $T_N x \ldots x T_N$ given by $(\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_N)$. The set $T_N x \ldots x T_N$ is denoted by T_N^N and the N(N-1) dimensional vector $(\underline{s}_1, \ldots, \underline{s}_N)$ is denoted by $\underline{t}$.

2.2 Transformations $\Pi_N \rightarrow T_N$ and $\Lambda^N \rightarrow T_N^N$

In this section the transformation that takes probability vectors to their representation in T_N is derived. Also, the transformation

that takes transition matrices in to their representation in T_N^N is derived.

First, the transformation from Π_N to T_N is specified. Vector $\underline{p} - \underline{\pi}_o$ lies in H if $\underline{p} \in \Pi_N$ and is given by

$$\underline{p} - \underline{\pi}_o = \sum_{i=1}^{N-1} s_i \underline{u}_i$$

The reciprocal basis $Y = \left\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_{N-1}\right\}$ of basis vectors χ is used to generate the expression

$$\langle \underline{p} - \underline{\pi}_o, \underline{v}_j \rangle = \left\langle \sum_{i=1}^{N-1} s_i \underline{u}_i, \underline{v}_j \right\rangle$$

$$= s_j \qquad j=1, \ldots, N-1$$

Therefore, the representation $\underline{s}$ of probability vector $\underline{p}$ satisfies

$$\underline{s} = (\underline{p} - \underline{\pi}_o) V$$

where the columns of matrix V are the reciprocal basis vectors.

Take any vector $\underline{s} \in T_N$. Vector $\underline{z} = \underline{s}U$ lies in H so there is a vector $\underline{p} \in \Pi_N$ such that $\underline{p} - \underline{\pi}_o = \underline{z}$ or

$$\underline{p} = \underline{\pi}_o + \underline{s}U$$

Next, the transformation from Λ^N to T_N^N is specified. Take any matrix $P \in \Lambda^N$. The rows of P lie in Π_N. Vector $\underline{s}_i$ is the unique representation of the i^{th} row $\underline{p}_i$ where

$$\underline{s}_i = (\underline{p}_i - \underline{\pi}_o)V \qquad i=1, \ldots, N$$

This expression specifies vector $\underline{t} = (\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_N)$, the representation of matrix P in set T_N^N.

Now, take any vector $\underline{t} \in T_N^N$. Vector $\underline{z}_i = \underline{s}_i U$ lies in H. There exists a vector $\underline{p}_i \in \Pi_N$ such that $\underline{p}_i = \underline{\pi}_o + \underline{s}_i U$. If the i^{th} row of an NxN matrix P is taken as $\underline{p}_i$, then matrix P lies in Λ^N.

2.3 Specifying a Basis χ

In this Section a particular basis χ is specified. This basis is used in many of the future developments and examples.

Let $\underline{e}_i$ be an N dimensional vector with all zero's except for the i^{th} component which is one. Since $\underline{\pi}_o$ can be any vector in Π_N define $\underline{\pi}_o$ as

$$\underline{\pi}_o = \underline{e}_1 \tag{2.5}$$

Define the basis χ by

$$\underline{u}_i = \frac{1}{\sqrt{2}}(\underline{e}_{i+1} - \underline{\pi}_o) \qquad i=1,\ldots,N-1 \tag{2.6}$$

In order that the vectors $\underline{u}_i$ defined above are basis vectors of H, each $\underline{u}_i$ must lies in H and the collection χ must be linearly independent. Vector $\underline{u}_i$ lies in H if $<\underline{u}_i, \underline{e}> = 0$. Substituting for $\underline{u}_i$ gives

$$\begin{aligned} <\underline{u}_i, \underline{e}> &= \frac{1}{\sqrt{2}} \left(<\underline{e}, \underline{e}_{i+1}> - <\underline{\pi}_o, \underline{e}>\right) \\ &= \frac{1}{\sqrt{2}} (1 - 1) \\ &= 0 \end{aligned}$$

To prove linear independence, take any vector $\underline{c} \in E^{N-1}$. If χ is a linearly independent collection then $\underline{c}U = \underline{0}$ implies that $\underline{c} = \underline{0}$. Vector $\underline{c}U$ is given by

$$\underline{c}U = \frac{1}{\sqrt{2}} \left[\sum_{i=1}^{N-1} c_i \underline{e}_{i+1} - \pi_o \sum_{i=1}^{N-1} c_i \right]$$

$$= \frac{1}{\sqrt{2}} \left[(0, c_1, c_2, \ldots, c_{N-1}) - (<\underline{c}_1 \underline{e}>, 0, \ldots, 0) \right]$$

$$= \frac{1}{\sqrt{2}} \left(-<\underline{c}_1 \underline{e}>, c_1, c_2, \ldots, c_{N-1} \right)$$

Clearly $\underline{c}U = 0$ implies that $\underline{c} = \underline{0}$ and hence χ is a linearly independent collection.

The reciprocal basis Y satisfies

$$<\underline{v}_j, \underline{u}_i> = \delta_{ij} \qquad i, j=1, \ldots, N-1$$

It can be easily shown that a reciprocal basis to χ is given by

$$\underline{v}_i = \sqrt{2}\, \underline{e}_{i+1} \qquad i=1, \ldots, N-1 \qquad (2.7)$$

<u>Example 2.1</u>: Consider a three state Markov chain. A basis for H, defined in Expression 2.6, is given by

$$\chi = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \ \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

The reciprocal basis, defined in Expression 2.7, is given by

$$Y = \left\{ \sqrt{2}\begin{pmatrix}0\\1\\0\end{pmatrix}, \sqrt{2}\begin{pmatrix}0\\0\\1\end{pmatrix} \right\}$$

Vectors $\underline{\pi}_o$, $\underline{u}_1$, $\underline{u}_2$, $\underline{v}_1$, and $\underline{v}_2$ are shown in Figure 2.5.

Set T_N in this basis is given by,

$$T_N = \left\{ \underline{s} \in E^{N-1} \mid \left(\frac{1}{\sqrt{2}} \sum_{i=1}^{N-1} s_i(\underline{e}_{i+1} - \underline{\pi}_o) + \underline{\pi}_o\right) \in \Pi_N \right\}$$

$$= \left\{ \underline{s} \in E^{N-1} \mid \left(\frac{1}{\sqrt{2}}(\sqrt{2} - \langle \underline{e}_1 \underline{s} \rangle)\underline{\pi}_o + \frac{1}{\sqrt{2}} \sum_{i=1}^{N-1} s_i \underline{e}_{i+1}\right) \in \Pi_N \right\}$$

or

$$T_N = \left\{ \underline{s} \in E^{N-1} \mid \left(\frac{1}{\sqrt{2}}(\sqrt{2} - \langle \underline{e}_1 \underline{s} \rangle), s_1/\sqrt{2}, \ldots, s_{N-1}/\sqrt{2}\right) \in \Pi_N \right\}$$

Since the above expression contains a probability vector, the components must satisfy the conditions

i) $0 \leq \frac{1}{\sqrt{2}}(\sqrt{2} - \langle e_1 \underline{s} \rangle) \leq 1$

ii) $0 \leq s_i/\sqrt{2} \leq 1 \qquad i = 1, \ldots, N-1$

Condition i) after some manipulation becomes

$$0 \leq \langle \underline{e}, \underline{s} \rangle \leq \sqrt{2}$$

Therefore, set T_N is defined as

$$T_N = \left\{ \underline{s} \in E^{N-1} \mid 0 \leq \langle \underline{e}, \underline{s} \rangle \leq \sqrt{2},\ 0 \leq s_i \leq \sqrt{2} \quad i=1,\ldots N-1 \right\}$$

Sets T_N for N = 2, 3, 4 are drawn in Figure 2.6.

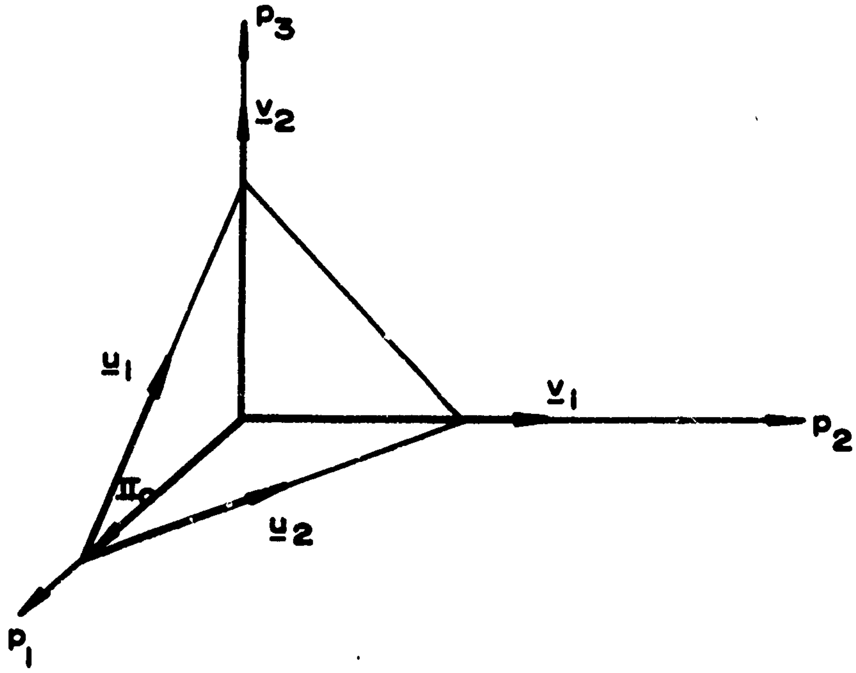


Figure 2.5 Sets χ and Y in E^3

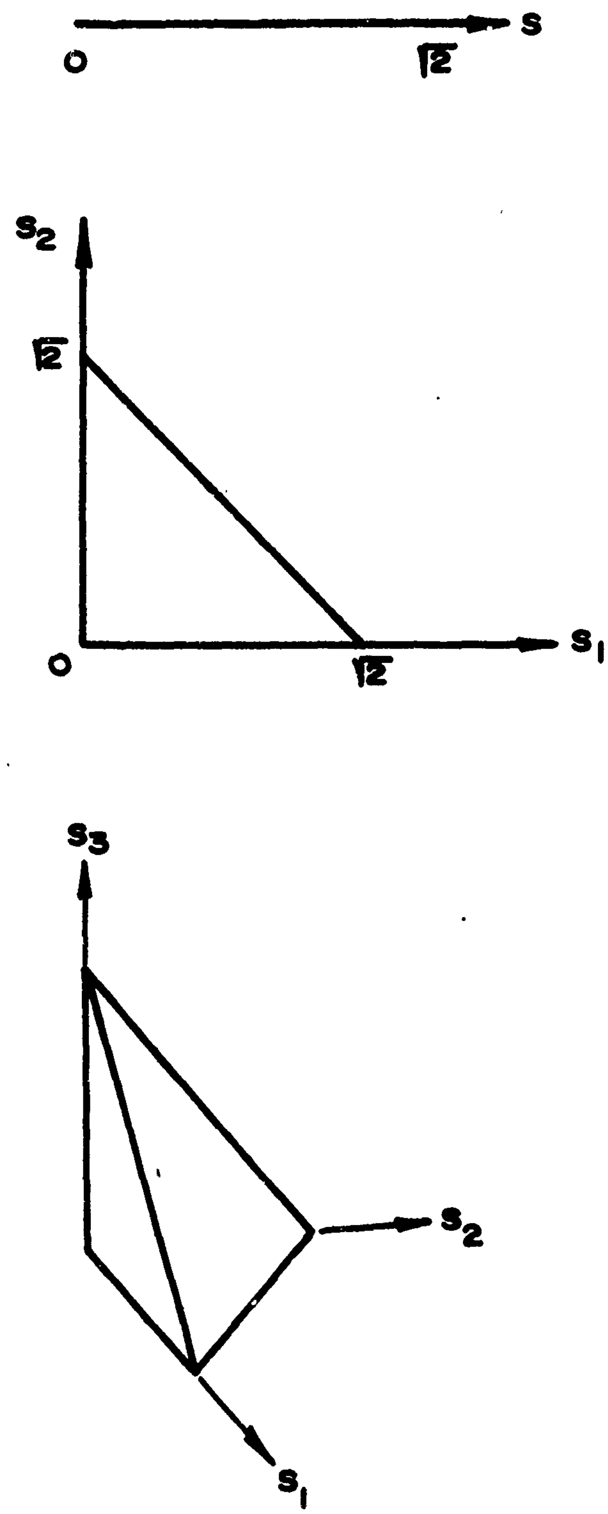


Figure 2.6 Sets T_2, T_3, and T_4

2.4 The Functions $\sigma:\Lambda^N \to \Pi_N$ and $w:T_N^N \to T_N$

The steady state probability vector $\underline{\pi}$ can be implicitly computed from matrix P by using Z-transforms or by computing the eigen vector of matrix P corresponding to the unity eigenvalue. In this section an explicit expression $\underline{\pi} = \underline{\sigma}(P)$ is derived and the desired expression $\underline{s} = \underline{w}(\underline{t})$ is specified.

2.4.1 The Function $\sigma:\Lambda^N \to \Pi_N$

A Markov chain with N states and ergodic transition matrix P is given. The probability distribution over the states at time t_n is given by the N-vector $\underline{\pi}(n)$. At time t_{n+1}, after one transition, the probability distribution over the state is

$$\underline{\pi}(n+1) = \underline{\pi}(n)\, P \tag{2.8}$$

Since P is ergodic $\underline{\pi}(n)$ converges to vector $\underline{\pi}$, the steady state probability vector, as $n \to \infty$.

Define NxN matrix P_o by

$$P_o = \begin{bmatrix} \underline{\pi}_o \\ \underline{\pi}_o \\ \vdots \\ \underline{\pi}_o \end{bmatrix} . \tag{2.9}$$

The rows of P_o are the vector $\underline{\pi}_o \in \Pi_N$. Given any vector $\underline{p} \in \Pi_N$ the following relationship holds,

$$\underline{p}\, P_o = \underline{\pi}_o$$

Therefore, $\underline{\pi}_o - \underline{p}\, P_o = \underline{0}$ for all $\underline{p} \in \Pi_N$.

The expression $\underline{\pi}_o - \underline{\pi}(n) P_o$ is added to Equation 2.8 to get

$$\underline{\pi}(n+1) = \underline{\pi}(n) P + \underline{\pi}_o - \underline{\pi}(n) P_o$$

Rearranging gives

$$\underline{\pi}(n+1) = \underline{\pi}(n) (P - P_o) + \underline{\pi}_o$$

The asymptotic behavior of this expression is

$$\underline{\pi} = \underline{\pi}(P - P_o) + \underline{\pi}_o$$

Solving for π gives

$$\underline{\pi} = \underline{\pi}_o (I - P + P_o)^{-1} \qquad (2.10)$$

Equation 2.10 specifies the mapping $\underline{\sigma}: P \rightarrow \underline{\pi}$.

The following Theorem proves the existence of $(I - P + P_o)^{-1}$ when P is ergodic.

Theorem 2.1: If P is ergodic then $(I - P + P_o)^{-1}$ exists.

Proof: The inverse of matrix $(I - P + P_o)$ exists if the null space, denoted by $N(I - P + P_o)$, is void. Take any $\underline{x} \in E^N$. Vector $\underline{x}$ lies in the null space if

$$\underline{x} (I - P + P_o) = \underline{0}$$

or equivalently

$$\underline{x} = \underline{x}\,P - \underline{x}\,P_o$$

$$= \sum_{i=1}^{N} x_i\,\underline{p}_i - \sum_{i=1}^{N} x_i\,\underline{\pi}_o$$

$$= \sum_{i=1}^{N} x_i\,(\underline{p}_i - \underline{\pi}_o)$$

Notice that $<\underline{p}_i - \underline{\pi}_o, \underline{e}> = \underline{0}$ for $i=1,\ldots,N$, so $\underline{p}_i - \underline{\pi}_o$ lies in the hyperplane H. A hyperplane is a linear subspace, so any linear combination of vectors in H will also lie in H. Therefore the null space is a subset of H.

Now, let $\underline{x}$ be any vector in H. Then $\underline{x}$ is written as

$$\underline{x} = \underline{\alpha}\,U$$

If $\underline{x} \in N(I - P + P_o)$ then

$$\underline{\alpha}\,U(I - P + P_o) = \underline{0}$$

or

$$\underline{\alpha}\,U - \underline{\alpha}\,UP + \underline{\alpha}\,UP_o = \underline{0} \qquad (2.11)$$

Matrix UP_o is evaluated in the following manner. The ij^{th} element is given by

$$\pi_{oj} < \underline{u}_i, \underline{e} > \qquad i,j=1,\ldots,N$$

Since $\underline{u}_i$ lies in H, $< \underline{u}_i, \underline{e} > = 0$. Therefore, UP_o is the zero matrix. Using this result, Equation 2.11 becomes

$$\underline{\alpha}\,U = \underline{\alpha}\,UP$$

or

$$\underline{x} = \underline{x}P \tag{2.12}$$

Equation 2.12 implies that vector $\underline{x} \epsilon H$ is an eigen vector associated with the unity eigen value. However, when P is ergodic, the steady state probability vector is the unique eigen vector associated with the unity eigen value [1]. Therefore, $\underline{x}$ is the zero vector and the null space of matrix $(I - P + P_o)$ is void.

QED

Example 2.3: Consider a two state Markov chain with transition matrix

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

The eigen vector $\underline{z}$ associated with the unity eigen value satisfies $\underline{z} = \underline{z}P$. Therefore,

$$\begin{aligned} \underline{z} &= (z_1, z_2) \\ &= (1/2\, z_1 + 1/4\, z_2, 1/2\, z_1 + 3/4\, z_2) \end{aligned}$$

or

$$2 z_1 = z_2$$

Since $\underline{z}$ is a probability vector

$$z_1 + z_2 = 3 z_1 = 1$$

and

$$\underline{z} = (1/3, 2/3)$$

Now, take $\underline{\pi}_o$ as (1, 0) then matrix $(I - P + P_o)$ is given by

$$\begin{pmatrix} 3/2 & -1/2 \\ 3/4 & 1/4 \end{pmatrix}$$

Substituting into Equation 2.10 gives

$$\underline{\pi} = (1, 0) \begin{pmatrix} 3/2 & -1/2 \\ 3/4 & 1/4 \end{pmatrix}^{-1}$$

$$= (1, 0) \begin{pmatrix} 1/3 & 2/3 \\ -1 & 2 \end{pmatrix}$$

$$= (1/3,\ 2/3)$$

2.4.2 <u>The Function $\underline{w}:\underline{t} \to \underline{s}$</u>

The equation $\underline{s} = \underline{w}(\underline{t})$ can be derived from equation $\underline{\pi} = \underline{\sigma}(P)$, using the transformations $\underline{\pi} \to \underline{s}$ and $P \to \underline{t}$. However, the equation $\underline{s} = \underline{w}(\underline{t})$ will be derived independent of $\underline{\pi} = \underline{\sigma}(P)$.

Vector $\underline{\pi}(n)$ can be expressed as

$$\underline{\pi}(n) = \underline{\pi}_o + \underline{s}(n)\, U$$

Substituting this expression into Equation 2.8 gives

$$\underline{s}(n+1)\, U = \underline{s}(n)\, UP + \underline{\pi}_o P - \underline{\pi}_o \tag{2.13}$$

Using the identity UV = I, Equation 2.13 is transformed to

$$\underline{s}(n+1) = \underline{s}(n)\, UPV + \underline{\pi}_o (PV - V) \tag{2.14}$$

The asymptotic behavior of Equation 2.14 is given by

$$\underline{s} = \underline{s}\, UPV + \underline{\pi}_o (PV - V) \tag{2.15}$$

Matrix PV can be developed further. The i^{th} row of matrix P is given by

$$\underline{p}_i = \underline{\pi}_o + \underline{s}_i U$$

Therefore, matrix P satisfies

$$P = P_o + SU \tag{2.16}$$

where matrix S is defined by

$$S = \begin{bmatrix} \underline{s}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{s}_N \end{bmatrix}$$

Post multiplying Expression 2.16 by V gives

$$PV = P_o V + S \tag{2.17}$$

since UV = I. Substitute Equation 2.17 into Equation 2.15 to get

$$\underline{s} = \underline{s}\, UP_o V + \underline{s}\, US + \underline{\pi}_o (P_o V - V + S)$$

Using the identity $UP_o = 0$ and the fact that $\underline{\pi}_o P_o = \underline{\pi}_o$,

$$\underline{s} = \underline{s}\, US + \underline{\pi}_o S$$

Solving for $\underline{s}$ gives

$$\underline{s} = \underline{\pi}_o\, S(I - US)^{-1} \tag{2.18}$$

Symbolically, Equation 2.18 is written as

$$\underline{s} = \underline{w}(\underline{t})$$

where

$$\underline{t} = (\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_N)$$

The following Theorem proves the existence of $(I - US)^{-1}$ when P is ergodic.

Theorem 2.2: If matrix P is ergodic, then $(I - US)^{-1}$ exists.

Proof: Matrix $I - US$ has an inverse if it has full rank. Expression $I - US$ can be written as

$$\begin{aligned} I - US &= I - UPV \\ &= UV - UPV + UP_oV \\ &= U(I - P + P_o)V \end{aligned}$$

Matrix $I - P + P_o$ has rank N when P is ergodic. Matrices U and V have rank N-1. Therefore, matrix $U(I - P + P_o)V$ is an N-1xN-1 matrix with rank N-1.

QED

Example 2.3: Consider a two state Markov chain. Take $\underline{\pi}_o$ and U as defined previously

$$\underline{\pi}_o = (1, 0)$$

$$\underline{u} = (-1, 1)$$

Equation 2.18 becomes

$$s = (1, 0) \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \left[1 - \frac{1}{\sqrt{2}}(-1, 1) \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \right]^{-1}$$

and

$$s = \frac{s_1 \sqrt{2}}{\sqrt{2} + s_1 - s_2} \tag{2.19}$$

2.5 Computing $f_\pi(\underline{s})$ from $f_P(\underline{t})$

The probability distribution over set Λ^N was given as

$$h_P(P) = h_1(\underline{p}_1)\ldots h_N(\underline{p}_N)$$

Density $h_P(P)$ can be easily transformed to density $f_P(\underline{t})$. The procedure will be outlined in the next Chapter. Density $f_P(\underline{t})$ is transformed to density $f_\pi(\underline{s})$ over the set T_N using the equation $\underline{s} = \underline{w}(\underline{t})$ in this section.

If the function $\underline{w}(\cdot)$ was one-to-one and continuous then the density $f_\pi(\underline{s})$ would be simply $f_P(\underline{w}^{-1}(\underline{s}))$. However, the function $\underline{w}(\cdot)$ is many-to-one and the inverse set $\underline{w}^{-1}(\underline{s})$ is dense in T_N^N. Therefore, density $f_\pi(\underline{s})$ is computed by first constructing the distribution function over set T_N and then taking the derivative with respect to vectors in T_N.

The distribution function, $F_\pi(\underline{a})$, is the probability mass over the set

$$A = \left\{\underline{s} \in T_N \,\middle|\, \underline{s} \leq \underline{a}\right\}$$

The inverse set $\underline{w}^{-1}(A)$ is given by

$$\underline{w}^{-1}(A) = \left\{\underline{t} \in T_N^N \mid \underline{w}(\underline{t}) \leq \underline{a}\right\}$$

The distribution function can be written as

$$F_\pi(\underline{a}) = \int_A f_\pi(\underline{s})\, d\underline{s}$$

or

$$F_\pi(\underline{a}) = \int_{\underline{w}^{-1}(A)} f_P(\underline{t})\, d\underline{t}$$

The density $f_\pi(\underline{s})$ is computed by taking the derivative of $F_\pi(\underline{a})$ at $\underline{a} = \underline{s}$,

$$f_\pi(\underline{s}) = \left. \frac{\partial^{N-1} F_\pi(\underline{a})}{\partial a_1 \ldots \partial a_{N-1}} \right|_{\underline{a} = \underline{s}} \qquad (2.20)$$

Consider the inverse set $\underline{w}^{-1}(A)$. The expression $\underline{w}(\underline{t}) \leq \underline{a}$ can be written as

$$\underline{\pi}_o S (I - US)^{-1} \leq \underline{a}$$

using Equation 2.18. Rearranging gives

$$(\underline{\pi}_o + \underline{a}U)S \leq \underline{a}$$

Let vector $\underline{\pi}_o + \underline{a}U$ be represented by vector $\underline{y}$, where

$$\underline{y} = \frac{1}{\sqrt{2}} (\sqrt{2} - \langle \underline{a}, \underline{e} \rangle, a_1, \ldots, a_{N-1})$$

if the basis χ defined previously is used. Notice that when $\underline{a} \in T_N$ then $\underline{y} \in \Pi_N$. Now, the expression $\underline{y}S$ is written as

$$\underline{y}S = \underline{t}\, B_a$$

where

$$B_a = \begin{bmatrix} \dfrac{\sqrt{2} - <\underline{a}_1\underline{e}>}{\sqrt{2}}\, I_{N-1} \\ \hdashline a_1/\sqrt{2} \quad I_{N-1} \\ \hdashline \vdots \\ \hdashline a_{N-1}/\sqrt{2}\, I_{N-1} \end{bmatrix}$$

and I_{N-1} is the (N-1)x(N-1) identity matrix. The notation B_a is employed to indicate that matrix B_a is a function of vector $\underline{a}$.

Using the above development, the inverse set becomes

$$\underline{w}^{-1}(A) = \left\{\underline{t} \in T_N^N \mid \underline{t}\, B_a \leq \underline{a}\right\} \tag{2.21}$$

Once vector $\underline{a}$ is selected, the above expression specifies a well defined subset of T_N^N. In the N state case (N>2) both integration and differentiation in T_N^N can be carried out on the computer.

Example 2.4: Consider a two state Markov chain with reward vector $\underline{r}$. The density $f_p(\underline{t}) = f_1(s_1)f_2(s_2)$ is assumed known. Set A is given by

$$A = \left\{\underline{s} \in T_2 \mid \underline{s} \leq \underline{a}\right\}$$

and set $\underline{w}^{-1}(A)$ is

$$\underline{w}^{-1}(A) = \left\{\underline{t} \in T_2^2 \mid \underline{t}\, B_a \leq \underline{a}\right\}$$

Expression $\underline{t}B_a$ becomes

$$\underline{t}B_a = (s_1, s_2)\begin{pmatrix} \dfrac{\sqrt{2} - a}{\sqrt{2}} \\ \dfrac{a}{\sqrt{2}} \end{pmatrix}$$

$$= \frac{s_1(\sqrt{2} - a) + s_2 a}{\sqrt{2}}$$

Therefore, set $\underline{w}^{-1}(A)$ is

$$\underline{w}^{-1}(A) = \left\{ \underline{t} \in T_2^2 \,\middle|\, s_1(\sqrt{2} - a) + s_2 a \leq \sqrt{2}\, a \right\}$$

and is drawn in Figure 2.7.

Density $f_\pi(\underline{a})$ is computed for two cases: $\frac{\sqrt{2}}{2} \leq a \leq \sqrt{2}$ and $0 \leq a \leq \sqrt{2}/2$

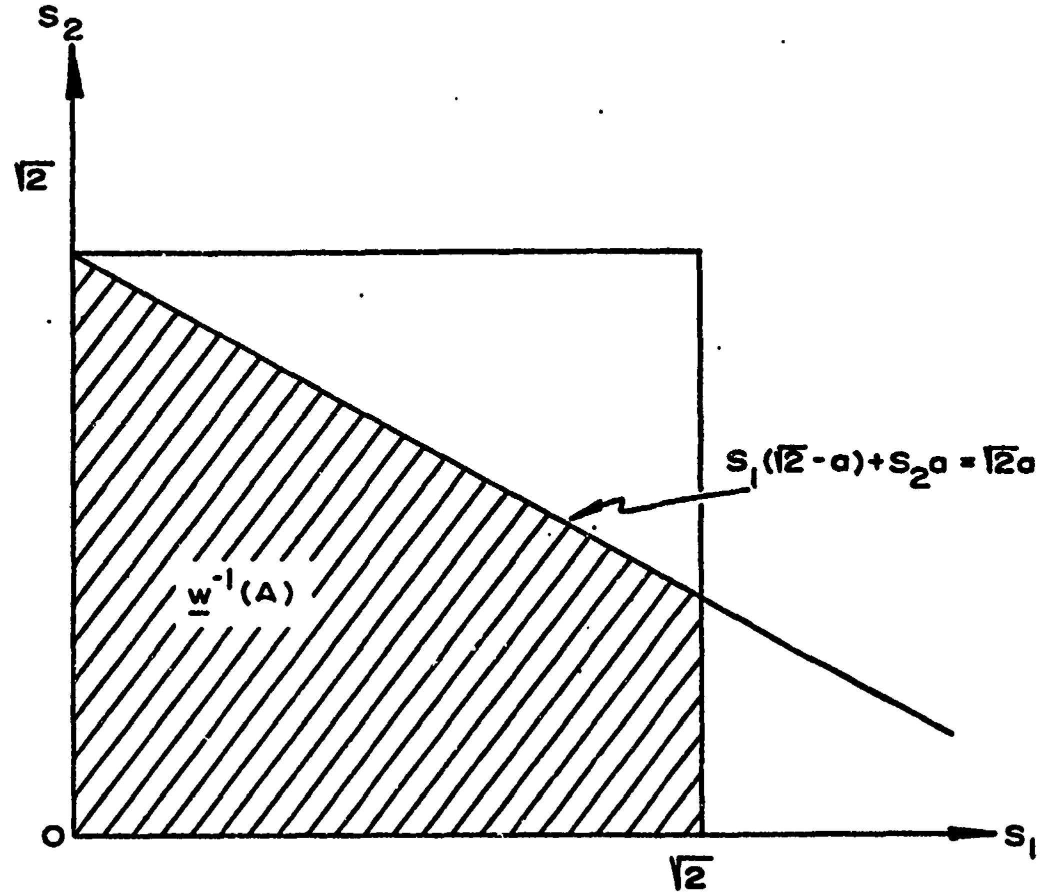


Figure 2.7 Set $\underline{w}^{-1}(A)$

CASE I: $\sqrt{2}/2 < a < \sqrt{2}$

$$F_\pi(A) = \int_{s_1=0}^{\sqrt{2}} \int_{s_2=0}^{\frac{a\sqrt{2}+as_1-\sqrt{2}s_1}{a}} f_1(s_1)\, f_2(s_2)\, ds_1 ds_2$$

using the fact that $f_P(s_1, s_2)$ factors into $f_1(s_1)f_2(s_2)$. Continuing,

$$F_\pi(A) = \int_{s_1=0}^{\sqrt{2}} f_1(s_1) \left\{ \int_{s_2=0}^{\frac{a\sqrt{2}+s_1(a-\sqrt{2})}{a}} f_2(s_2) ds_2 \right\} ds_1$$

$$= \int_{s_1=0}^{\sqrt{2}} f_1(s_1)\, g(s_1, a)\, ds_1$$

Computing $f_\pi(\underline{a})$ from Equation 2.20 gives

$$f_\pi(a) = \frac{\partial}{\partial a} \int_{s_1=0}^{\sqrt{2}} f_1(s_1)\, g(s_1, a)\, ds_1$$

$$= \int_{s_1=0}^{\sqrt{2}} f_1(s_1) \frac{\partial}{\partial a}\, g(s_1, a)\, ds_1 \qquad (2.22)$$

The function $\frac{\partial}{\partial a} g(s_1, a)$ is evaluated using Leibenitz's rule,

$$\frac{\partial}{\partial a} g(s_1, a) = f_2\left(\frac{a\sqrt{2}+s_1(\sqrt{2}-a)}{a}\right) \frac{\partial}{\partial a}\left[\frac{a\sqrt{2}+s_1(\sqrt{2}-a)}{a}\right]$$

$$= f_2\left(\frac{a\sqrt{2}+s_1(\sqrt{2}-a)}{a}\right) \left(\frac{s_1\sqrt{2}}{a^2}\right) \qquad (2.23)$$

Substituting Equation 2.23 into Equation 2.22 gives

$$f_\pi(a) = \int_{s_1=0}^{\sqrt{2}} f_1(s_1)\, f_2\left(\frac{a\sqrt{2}-s_1(\sqrt{2}-a)}{a}\right) \frac{s_1\sqrt{2}}{a^2}\, ds_1$$

<u>CASE II</u>: $0 \le a \le \sqrt{2}/2$

$$F_\pi(A) = \int_{s_2=0}^{\sqrt{2}} f_2(s_2) \left\{ \int_{s_1=0}^{\frac{a(\sqrt{2}-s_2)}{(\sqrt{2}-a)}} f_1(s_1)\, ds_1 \right\} ds_2$$

$$= \int_{s_2=0}^{\sqrt{2}} f_2(s_2)\, h(s_2, a)\, ds_2$$

$$f_\pi(a) = \int_{s_2=0}^{\sqrt{2}} f_2(s_2) \frac{\partial}{\partial a} h(s_2, a)\, ds_2$$

Finally,

$$f_\pi(a) = \int_{s_2=0}^{\sqrt{2}} f_2(s_2)\, f_1\left(\frac{a(\sqrt{2}-s_2)}{\sqrt{2}-a}\right) \frac{2-\sqrt{2}\, s_2}{(\sqrt{2}-a)^2}\, ds_2$$

To summarize,

$$f_\pi(s) = \begin{cases} \displaystyle\int_{s_2=0}^{\sqrt{2}} f_2(s_2) f_1\left(\frac{s(\sqrt{2}-s_2)}{\sqrt{2}-a}\right) \frac{2-\sqrt{2}s_2}{(\sqrt{2}-a)^2}\, ds_2, \quad 0 \le s \le \frac{\sqrt{2}}{2} & (2.24) \\ \displaystyle\int_{s_1=0}^{\sqrt{2}} f_1(s_1) f_2\left(\frac{a\sqrt{2}-s_1(\sqrt{2}-a)}{a}\right) \frac{s_1\sqrt{2}}{a^2}\, ds_1, \quad \frac{\sqrt{2}}{2} \le s \le \sqrt{2} & (2.25) \end{cases}$$

2.6 Evaluating $E(\Delta)$

The gain is given by

$$\Delta = \langle \underline{\pi}, \underline{r} \rangle \tag{2.26}$$

Probability vector $\underline{\pi}$ can be written as

$$\underline{\pi} = \underline{\pi}_o + \underline{s}U$$

or

$$\underline{\pi} = \underline{\pi}_o + \underline{w}(\underline{t})\, U$$

Substituting into Equation 2.26 gives

$$\Delta = \langle \underline{\pi}_o, \underline{r} \rangle + \langle \underline{w}(\underline{t})U, \underline{r} \rangle$$

Since density $h_P(P)$ is assumed known, density $f_P(\underline{t})$ is known. Therefore, $E(\Delta)$ becomes,

$$E(\Delta) = \int_{T_N^N} \langle \underline{w}(\underline{t})U, \underline{r} \rangle f_P(\underline{t})\, d\underline{t} + \langle \underline{\pi}_o, r \rangle \tag{2.27}$$

where

$$\underline{w}(\underline{t}) = \underline{\pi}_o S(I - US)^{-1}$$

Example 2.5: Consider a two state Markov chain with known density $f_P(\underline{t}) = f_1(s_1)f_2(s_2)$. The reward vector is $\underline{r}$. The expected value of the gain using Equation 2.27 and Equation 2.19 is

$$E(\Delta) = \langle \underline{u}, \underline{r} \rangle \int_{T_2^2} \frac{s_1\sqrt{2}}{\sqrt{2}+s_1-s_2} f_1(s_1)f_2(s_2)ds_1ds_2 + \langle \underline{\pi}_o, \underline{r} \rangle$$

2.7 Ergodicity

The concept of ergodicity in the set T_N^N is developed next. If a transition matrix has all non-zero entries, then it is ergodic [9]. Since the equation $\underline{\pi} = \underline{\sigma}(P)$ holds for all ergodic matrices P, the equation $\underline{s} = \underline{w}(\underline{t})$ holds for all vectors $\underline{t}$ that represent ergodic matrices. The subset of T_N^N that corresponds to non-ergodic matrices must be identified. If this subset is assigned non-zero probability mass by the matrix beta density, the technique of computing $f_\pi(\underline{s})$ from $f_P(\underline{t})$ using $\underline{s} = \underline{w}(\underline{t})$ may not be valid.

Take a vector $\underline{p} \in \Pi_N$. Probability vector $\underline{p}$ can be thought of as a row of matrix P and is written as

$$\underline{p} = \underline{\pi}_o + \underline{s}\, U$$

$$= (1, 0, \ldots, 0) + \underline{s}\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 & 0 \ldots\ldots 0 \\ -1 & 0 & 1 \;\; 0 \ldots 0 \\ & & \vdots \\ -1 & 0 & \ldots\ldots 01 \end{bmatrix}$$

$$= (1, 0, \ldots, 0) + \frac{1}{\sqrt{2}}(- <\underline{s},\underline{e}> s_1, \ldots, s_N\text{-}1)$$

$$= (1 - \frac{1}{\sqrt{2}} <\underline{s}, \underline{e}>, s_1, \ldots, s_{N-1})$$

If $p_i = 0$ then either $s_{i-1} = 0$ or $<\underline{s}, \underline{e}> = \sqrt{2}$, in the case where i=1. From the definition of set T_N^N in set χ, these two conditions place the resulting vector $\underline{t}$ on the boundary of T_N^N.

In the next Chapter the matrix beta density is defined. This density assigns all its probability mass to transition matrices that have all non-zero entries. Therefore, the matrix beta density

assigns zero probability to the boundary of T_N^N. Recall that equation $\underline{s} = \underline{w}(\underline{t})$ does not hold on the boundary of T_N^N. Therefore, the integral of $\underline{w}(\underline{t})$ times the matrix beta density over T_N^N exists. Further, the method of computing $f_\pi(\underline{s})$ from $f_p(\underline{t})$ using $\underline{s} = \underline{w}(\underline{t})$ is valid

Chapter 3

COMPUTING THE EXPECTED VALUE OF THE GAIN CONDITIONED ON OBSERVATIONS

In the last Chapter the uncertainty over the transition probabilities was characterized by a general density function $h_{\underline{P}}(P)$. In this Chapter density $h_{\underline{P}}(P)$ is specified as matrix beta. Observations $\underline{x}_n$ are recorded from the states of nature Ω . The posteriori density is computed using Bayes' formula. The expected value of the gain conditioned on the observations is computed from the posteriori density.

3.1 The Matrix Beta Density

The transition probabilities $P = [p_{ij}]$ are said to have the matrix beta density with "parameter" $M = [m_{ij}]$ if P has the joint density

$$f^N_{M\beta}(P|M) = \begin{cases} k(M) \prod\limits_{\substack{k=1 \\ j=1}}^{N} (p_{ij})^{m_{ij}-1} , & P \in \Lambda^N \\ 0 & , \text{ elsewhere} \end{cases}$$

The normalizing constant k(M) is given by

$$k(M) = \prod_{i=1}^{N} \frac{\Gamma(M_i)}{\prod\limits_{j=1}^{N} \Gamma(m_{ij})}$$

where

$$M_i = \sum_{j=1}^{N} m_{ij} \qquad i=1,\ldots,N$$

The "parameter" M is an NxN matrix such that $m_{ij} > 0 \quad i,j=1,\ldots,N$.

Notice that the rows of $\underline{P}$ are independent random probability vectors when P has the matrix beta density. Therefore, the density function can be factored as

$$f_{M\beta}^{N}(P|M) = \prod_{i=1}^{N} h_i(\underline{p}_i|\underline{m}_i)$$

where $\underline{m}_i$ is the i^{th} row of matrix M. The probability density $h_i(\underline{p}_i|\underline{m}_i)$ is given by

$$h_i(\underline{p}_i|\underline{m}_i) = k(\underline{m}_i) \prod_{j=1}^{N} (p_{ij})^{m_{ij}-1} \tag{3.1}$$

where

$$k(\underline{m}_i) = \frac{\Gamma(M_i)}{\prod_{j=1}^{N} \Gamma(m_{ij})}$$

Density $h_i(\underline{p}_i|\underline{m}_i)$ is the density over the i^{th} row of matrix P, and is called the vector beta density with "parameter" $\underline{m}_i$.

The following statistics on $\tilde{p}_{ij}$ are derived using the density function in Expression 3.1:

$$\text{i)}\quad E(\tilde{p}_{ij}) = \bar{p}_{ij} = \frac{m_{ij}}{M_i}$$

$$\text{ii)}\quad Var(\tilde{p}_{ij}) = \frac{m_{ij}(M_i - m_{ij})}{M_i^2(M_i+1)} \tag{3.2}$$

$$\text{iii)}\quad Cov(\tilde{p}_{ij}, \tilde{p}_{in}) = \frac{-m_{ij}\, m_{in}}{M_i^2(M_i+1)}$$

3.2 The A Priori Density

The states of nature Ω in the case of a single Markov chain is just the set Λ^N. The a priori density over Λ^N is specified as the matrix beta density with "parameter" M,

$$\xi_o(\omega) = f^N_{M\beta}(P|M)$$

Given this definition, the problem becomes how to specify the elements of matrix M so that the a priori density reflects, in some way, the a priori knowledge of the transition probabilities. Since the transition vectors $\tilde{\underline{p}}_i$ are probabilistically independent, the rows of M can be selected independently.

For example, suppose the value of $\underline{m}_i$ is specified as

$$m_{i1} = m_{i2} = \ldots = m_{iN} = m$$

The expected value of $\tilde{p}_{ij}$ is given by condition i) of Equation 3.2,

$$E(\tilde{p}_{ij}) = \frac{m_{ij}}{M_i}$$

$$= \frac{m}{Nm}$$

$$= \frac{1}{N} \qquad j=1,\ldots,N$$

The magnitude of m does not affect the expected value. However, the magnitude does affect the variance and covariance. Substituting $m_{ij} = m$ into condition ii) and iii) gives

$$\mathrm{Var}(\tilde{p}_{ij}) = \frac{N-1}{N^2\,(Nm+1)} \qquad j=1,\ldots,N$$

$$\mathrm{Cov}(\tilde{p}_{ij}\,\tilde{p}_{in}) = \frac{-1}{N^2(Nm+1)} \qquad j=1,\ldots,N$$

As m increases $\mathrm{var}(p_{ij})$ and $\mathrm{cov}(p_{ij}p_{in})$ go to zero.

The example illustrates the fact that the relative proportions of the m_{ij} will specify the expected value of the $\tilde{p}_{ij}$ and the magnitude will specify the variance and covariance. In otherwords, if $\bar{\underline{p}}_i$ is the expected value of $\tilde{\underline{p}}_i$ then $\underline{m}_i$ will be given by $\underline{m}_i = \alpha\,\bar{\underline{p}}_i$ where $\alpha > 0$. The magnitude of α specifies the variance and covariance.

<u>Example 3.1</u>: Consider a two state Markov chain. Let

$$m_{i1} = m_{i2} = k \qquad i=1,2$$

where k is a positive integer. The expected value of $\tilde{p}_{ij}$ is given by

$$E(\tilde{p}_{ij}) = 1/2 \qquad i,j=1,2$$

The variance and covariance of $\tilde{p}_{ij}$ are given by

$$\mathrm{Var}(\tilde{p}_{ij}) = \frac{1}{4(2k+1)}$$

$$\mathrm{Cov}(\tilde{p}_{ij}\ \tilde{p}_{in}) = \frac{-1}{4(2k+1)}$$

The values of $\mathrm{var}(\tilde{p}_{ij})$ and $\mathrm{cov}(\tilde{p}_{ij}\ \tilde{p}_{in})$ for several values of k are listed in Table 3.1.

For each value of k the densities $h_i(\underline{p}_i|\underline{m}_i)$ i=1, 2 are given by

$$h_i(\underline{p}_i|\underline{m}_i) = k(\underline{m}_i) \prod_{j=1}^{N} (p_{ij})^{k-1}$$

$$k(\underline{m}_i) = \frac{\Gamma(2k)}{\Gamma(k)\Gamma(k)}$$

Density $h_i(\underline{p}_i|\underline{m}_i)$ can be drawn as in Figure 2.1, or can be transformed to a density $f_i(s_i|\underline{m}_i)$ as drawn in Figure 2.2. The details on how $f_i(s_i|\underline{m}_i)$ is computed is covered in Section 3.4. However, as seen from Figures 2.1 and 2.2 density $f_i(s_i|\underline{m}_i)$ has the same shape as density $h_i(\underline{p}_i|\underline{m}_i)$ in the two state case, and therefore should cause no confusion. The transformed density $f_i(s_i|\underline{m}_i)$ is drawn in Figure 3.1 for several values of k. The density is seen to concentrate more of its probability mass near s=0.707 as k increases. This is equivalent to saying that the covariance and variance are going to zero as k increases.

The following method is proposed to choose matrix M in the a priori density.

Table 3.1

k	$Var(p_{ij})$	$Cov(p_{ij}\ p_{in})$
1	0.05	-0.05
2	0.028	-0.028
5	0.012	-0.012
15	0.004	-0.004
30	0.002	-0.002

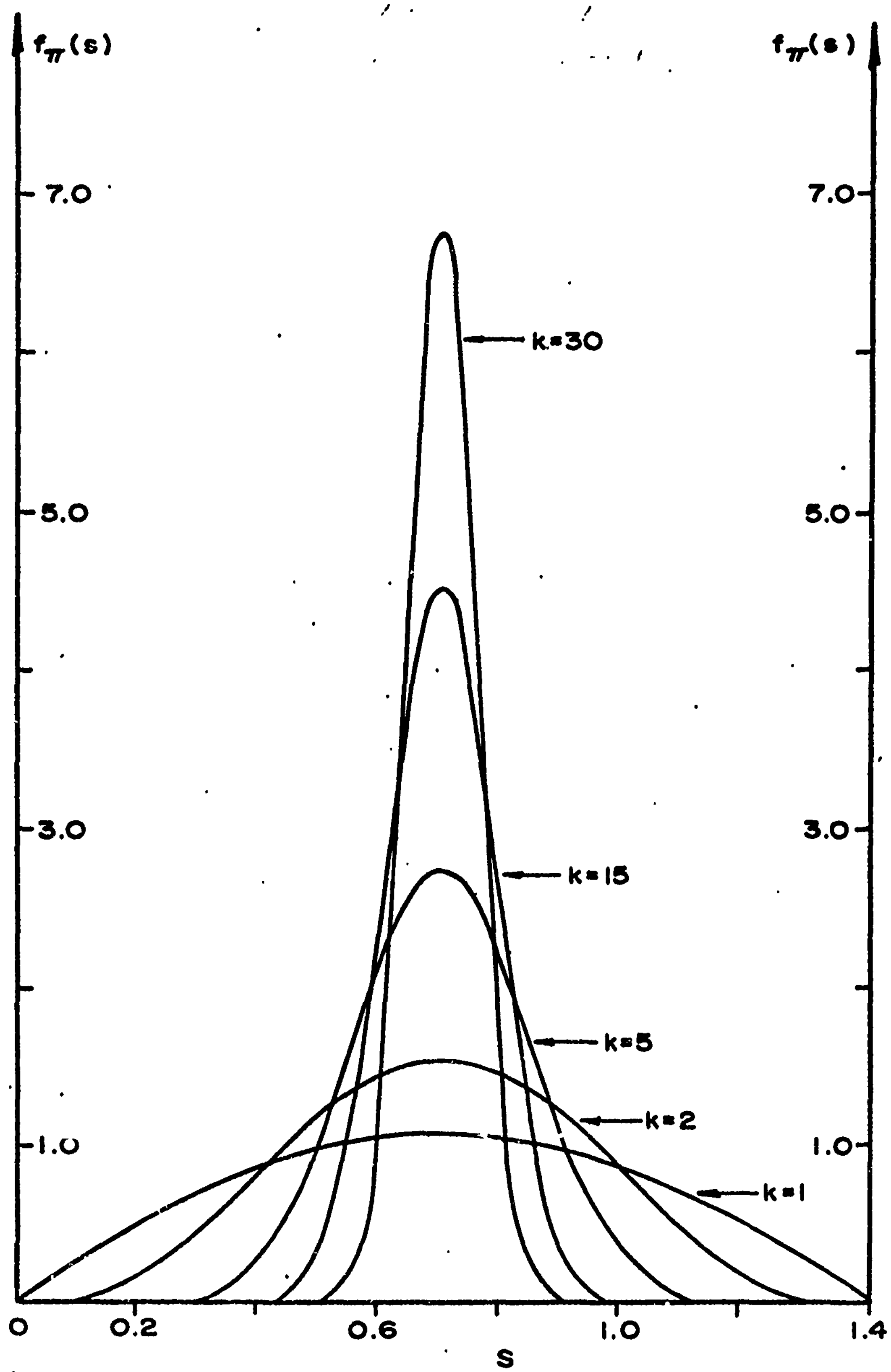


Figure 3.1 A Priori Density

1. If the transition probabilities are completely unknown set

$$m_{ij} = 1 \qquad i,j=1,\ldots,N$$

2. If the knowledge of the transition probabilities is more precise, select vectors $\underline{m}_i$ $i=1,\ldots,N$ such that the resulting expected values $E(\tilde{p}_{ij})$ lie in the expected range. Select the magnitude to be proportional, in some way, to this knowledge.

3.3 The Posteriori Density

The posteriori density over Λ^N is computed from the a priori density, observations taken from the Markov chain, and Bayes' formula. Let $\underline{x}_n = \{x_o, x_1, \ldots, x_n\}$ be a sequence of states observed from the N state Markov chain where x_o, the initial state, is known. The probability of observing $\underline{x}_n$ is given by $p_{x_o x_1} p_{x_1 x_2} \ldots p_{x_{n-1} x_n}$, where $p_{x_i x_{i+1}}$ is the probability of making a transition from state i to state j in sequence $\underline{x}_n$. Let f_{ij} denote the number of transitions observed from state i to state j in sequence $\underline{x}_n$. The NxN matrix $F = [f_{ij}]$ is called the transition count matrix of the sample $\underline{x}_n$.

The conditional probability of observing $\underline{x}_n$ given a $P \in \Lambda^N$ is denoted by $l(\underline{x}_n | P)$ where

$$l(\underline{x}_n | P) = p_{x_o x_1} \ldots p_{x_{n-1} x_n}$$
$$= \prod_{\substack{i=1 \\ j=1}}^{N} (p_{ij})^{f_{ij}}$$

The function $l(\underline{x}_n | P)$ is called the likelihood function.

The posteriori density over Λ^N, denoted by $\xi(P | \underline{x}_n)$, is computed using Bayes' formula

$$\xi(P | \underline{x}_n) = \frac{l(\underline{x}_n | P)\, \xi_o(P)}{\int_{\Lambda^N} l(\underline{x}_n | P)\, \xi_o(P)\, dP}$$

$$= \frac{1}{\int_{\Lambda^N} l(\underline{x}_n | P)\, \xi_o(P)\, dP} \prod_{\substack{i=1 \\ i=j}}^{N} (p_{ij})^{m_{ij}+f_{ij}-1}$$

Thus, $\xi(P|\underline{x}_n)$ has the same form as the matrix beta density with parameter M+F. Since $\int_{\Lambda^N} \xi(P|\underline{x}_n)dP=1$, $\xi(P|\underline{x}_n)$ has the form

$$\xi(P|\underline{x}_n) = f^N_{M\beta}(P|M+F) \qquad (3.3)$$

Martin [9] proved that as the number of observations goes to infinity the density $f^N_{M\beta}(P|M+F)$ will concentrate on one $P \in \Lambda^N$.

3.4 The Posteriori Density $f_P(\underline{t}|M+F)$

This section deals with the transformation of $f^N_{M\beta}(P|M+F)$ to $f_P(\underline{t}|M+F)$.

Given the matrix beta density over Λ^N,

$$f^N_{M\beta}(P|M+F) = h_1(\underline{p}_1|\underline{m}_1+\underline{f}_1)\ldots h_N(\underline{p}_N|\underline{m}_N+\underline{f}_N)$$

the corresponding density over T^N_N

$$f_P(\underline{t}|M+F) = f_1(\underline{s}_1|\underline{m}_1+\underline{f}_1)\ldots f_N(\underline{s}_N|\underline{m}_N+\underline{f}_N)$$

is derived by transforming the factored density $h_i(\underline{p}_i|\underline{m}_i+\underline{f}_i)$ to the factored density $f_i(\underline{s}_i|\underline{m}_i+\underline{f}_i)$ instead of transforming $f^N_{M\beta}(P|M+F)$ to $f_P(\underline{t}|M+F)$ directly. A basis for E^N is given by $\chi' = \{\chi, \pi_o\}$. Transition probability vectors $\underline{p}_i$ have the following representation in basis χ'

$$(s_{i1}, s_{i2}, \ldots, s_{iN-1}, 1) \qquad i=1,\ldots,N$$

The transformation that takes N vectors in the natural basis to N vectors in basis χ' is the NxN matrix V'

$$V' = \left[V \vdots \underline{e}^T \right]$$

Therefore,

$$(s_{i1}, \ldots, s_{iN-1}, 1) = (p_{i1}, \ldots, p_{iN})\, V'$$
$$= (<\underline{p}_i, \underline{v}_1>, \ldots, <\underline{p}_i, \underline{v}_{N-1}>, <\underline{p}_i, \underline{e}>)$$

The Jacobian of this transformation is the NxN matrix J where

$$J = \begin{bmatrix} V^T \\ \cdots\cdot \\ \underline{e} \end{bmatrix}$$

The density $f_i(\underline{s}_i | \underline{m}_i + \underline{f}_i)$ is given by

$$f_i(\underline{s}_i | \underline{m}_i + \underline{f}_i) = \frac{1}{\det(J)}\; h_i(\underline{\pi}_o + \underline{s}_i U | \underline{m}_i + \underline{f}_i) \tag{3.4}$$

Therefore

$$f_P(\underline{t} | M+F) = \prod_{i=1}^{N} \frac{1}{\det(J)}\; h_i(\underline{\pi}_o + \underline{s}_i U | \underline{m}_i + \underline{f}_i) \tag{3.5}$$

Example 3.2: Consider a two state Markov chain. Density $f^2_{M\beta}(P|M+F)$ is given by

$$f^2_{M\beta}(P|M+F) = h_1(\underline{p}_i | \underline{m}_1 + \underline{f}_1)\; h_2(\underline{p}_2 | \underline{m}_2 + \underline{f}_2)$$

where

$$h_i(\underline{p}_i | \underline{m}_i + \underline{f}_i) = \frac{\Gamma(M_i+F_i)}{\Gamma(m_{i1}+f_{i1})\Gamma(m_{i2}+f_{i2})} \; p_{i1}^{m_{i1}+fi_1} \; p_{i2}^{m_{i2}+f_{i2}} \quad i=1,2$$

The density over T_2^2 $f_P(\underline{t} | M+F)$ is given by

$$f_P(\underline{t} | M+F) = f_1(\underline{s}_1 | \underline{m}_1+\underline{f}_1)\, f_2(\underline{s}_2 | \underline{m}_2+\underline{f}_2)$$

where

$$f_i(\underline{s}_i | \underline{m}_i+\underline{f}_i) = \frac{\Gamma(M_i+F_i)}{\sqrt{2}\,\Gamma(m_{i1}+f_{i1})\Gamma(m_{i2}+f_{i2})} \left(1 - \frac{s_i}{\sqrt{2}}\right)^{m_{i1}+f_{i1}} \times$$

$$\left(\frac{s_i}{\sqrt{2}}\right)^{m_{i2}+f_{i2}} \quad i=1,2 \qquad (3.6)$$

and

$$\det[J] = \det \begin{bmatrix} 1 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$= \sqrt{2}$$

3.5 The Posteriori Density $f_\pi(\underline{s} | M+F)$

The posteriori density over T_N $f_\pi(\underline{s} | M+F)$ is computed from the posteriori density over T_N^N $f_P(\underline{t} | M+F)$ using the results of Chapter 2. Density $f_\pi(\underline{s} | M+F)$ is given by

$$f_\pi(\underline{s} | M+F) = \left. \frac{\partial^{N-1} F_\pi(\underline{a})}{\partial a_1 \cdots \partial a_{N-1}} \right|_{\underline{a}=\underline{s}}$$

where

$$F_\pi(\underline{a}) = \int_{\underline{w}^{-1}(A)} f_P(\underline{t} \mid M+F)\, d\underline{t}$$

$$\underline{w}^{-1}(A) = \left\{ \underline{t} \in T_N^N \mid \underline{t}\, B_a \leq \underline{a} \right\}$$

Consider the special case of a two state Markov chain. Density $f_\pi(\underline{s} \mid M)$ is computed by substituting Expression 3.6 with $F = [0]$ into Equations 2.24 and 2.25,

$$f_\pi(s \mid M) = \begin{cases} k'(M) \displaystyle\int_{s_2=0}^{\sqrt{2}} (\sqrt{2}-s_2)^{m_{21}-1} s_2^{m_{22}-1} \left(\frac{2-2\sqrt{2}s+ss_2}{\sqrt{2}-s}\right)^{m_{11}-1} \\ \quad \times \left(\dfrac{s(\sqrt{2}-s_2)}{\sqrt{2}-s}\right)^{m_{12}-1} \left(\dfrac{2-\sqrt{2}\, s_2}{(\sqrt{2}-s)^2}\right) ds_2, \quad 0 < s \leq \sqrt{2}/2 \\ k'(M) \displaystyle\int_{s_1=0}^{\sqrt{2}} (\sqrt{2}-s_1)^{m_{11}-1} s_1^{m_{12}} \left(\frac{s_1(\sqrt{2}-s)}{s}\right)^{m_{21}-1} \\ \quad \times \left(\dfrac{s\sqrt{2}-s_1(\sqrt{2}-s)}{s}\right)^{m_{22}-1} \left(\dfrac{\sqrt{2}}{s^2}\right) ds_1 \quad \dfrac{\sqrt{2}}{2} < s \leq \sqrt{2} \end{cases}$$

where

$$k'(M) = k(M)\, \sqrt{2}^{\,4-M_1-M_2}$$

The above expression can be manipulated to get

$$f_\pi(s|M) = \begin{cases} \frac{1}{\sqrt{2}} k'(M) \left(\frac{s}{\sqrt{2}-s}\right)^{m_{12}} \left(\frac{\sqrt{2}}{\sqrt{2}-s}\right)^2 \sum_{j=0}^{m_{11}-1} \xi_j \beta_j(s), & 0<s\leq \sqrt{2}/2 \\ k'(M) \left(\frac{\sqrt{2}-s}{s}\right)^{m_{22}+m_{21}-2} \left(\frac{\sqrt{2}}{s^2}\right) \sum_{j=0}^{m_{22}-1} x_j(s) w_j, & \frac{\sqrt{2}}{2} < s \leq \sqrt{2} \end{cases} \tag{3.7}$$

where

$$\xi_j = \sum_{i=0}^{m_{21}+m_{12}-1} \alpha_i \gamma_{ij}$$

$$\gamma_{ij} = \frac{\sqrt{2}^{\,i+j+m_{22}}}{i+j+m_{22}}$$

$$\sum_{i=0}^{m_{12}+m_{21}-1} \alpha_i s_2^i = (\sqrt{2} - s_2)^{m_{21}+m_{12}-1}$$

$$\sum_{j=0}^{m_{11}-1} \beta_j(s)\, s_2^j = \left(\frac{2-2\sqrt{2}\, s}{s} + s_2\right)^{m_{11}-1}$$

$$w_j = \sum_{i=0}^{m_{11}-1} \varphi_{ij}\, \tau_i$$

$$\varphi_{ij} = \frac{\sqrt{2}^{\,i+j+m_{21}+m_{12}}}{i+j+m_{21}+m_{12}}$$

$$\sum_{i=0}^{m_{11}-1} \tau_i s_1^i = (\sqrt{2} - s_1)^{m_{11}-1}$$

$$\sum_{j=0}^{m_{22}-1} x_j(s)\, s_1^j = \left(\frac{s\sqrt{2}}{\sqrt{2} - s} - s_1\right)^{m_{22}-1}$$

Example 3.3: Consider a two state Markov chain. The a priori density is matrix beta with parameter M given by

$$M = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

Desnity $f_\pi(s|M)$ as specified in Expression 3.7 is

$$f_\pi(s|M) = \begin{cases} 9\left(\frac{s}{\sqrt{2}-s}\right)\left(\frac{\sqrt{2}}{\sqrt{2}-s}\right)^2 \left[\frac{2\sqrt{2}}{5}\left(\frac{\sqrt{2}-2s}{s}\right)^2 + \frac{152}{15}\left(\frac{1-\sqrt{2}\,s}{s}\right) + \frac{2\sqrt{2}}{35}\right], & 0 < s \le \frac{\sqrt{2}}{2} \\ 9\sqrt{2}\left(\frac{\sqrt{2}-s}{s}\right)^2\left(\frac{\sqrt{2}}{s}\right)\left[\frac{2}{15}\,\frac{s\sqrt{2}}{\sqrt{2}-s} - \frac{8\sqrt{2}}{105}\right], & \frac{\sqrt{2}}{2} < s \le \sqrt{2} \end{cases}$$

Density $f_\pi(s|M)$ is drawn in Figure 3.2.

Example 3.4: Consider a two state Markov chain with transition matrix

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The steady state probability vector $\underline{\pi}$ is

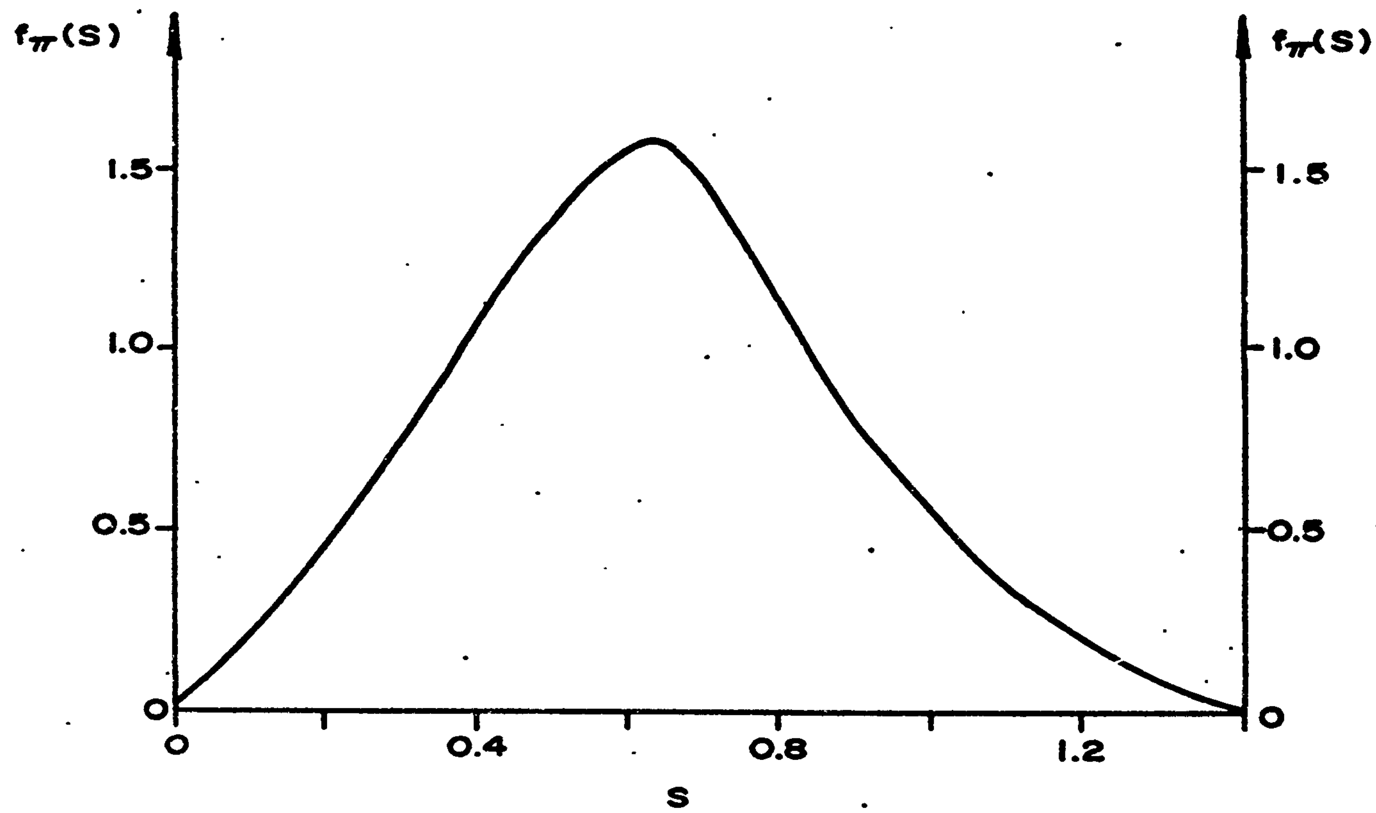


Figure 3.2 A Priori Density

$$\underline{\pi} = (5/7 \quad 5/7)$$

and its representation in T_2 is the scalar s,

$$s = 0.404$$

The a priori density over P is defined as matrix beta with parameter M,

$$M = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

The transition count matrix F was computed from observations taken from a computer simulation of the above defined Markov chain. Matrix F at 100 transitions was

$$F = \begin{bmatrix} 74 & 16 \\ 16 & 19 \end{bmatrix}$$

and at 250 transitions it was

$$F = \begin{bmatrix} 141 & 35 \\ 34 & 40 \end{bmatrix}$$

The posteriori density $f_\pi(s|M+F)$ was computed by numerically integrating Equation 3.7 with M replaced by M+F. The resulting densities are drawn in Figure 3.3. The posteriori density is indeed concentrating on $s = 0.404$.

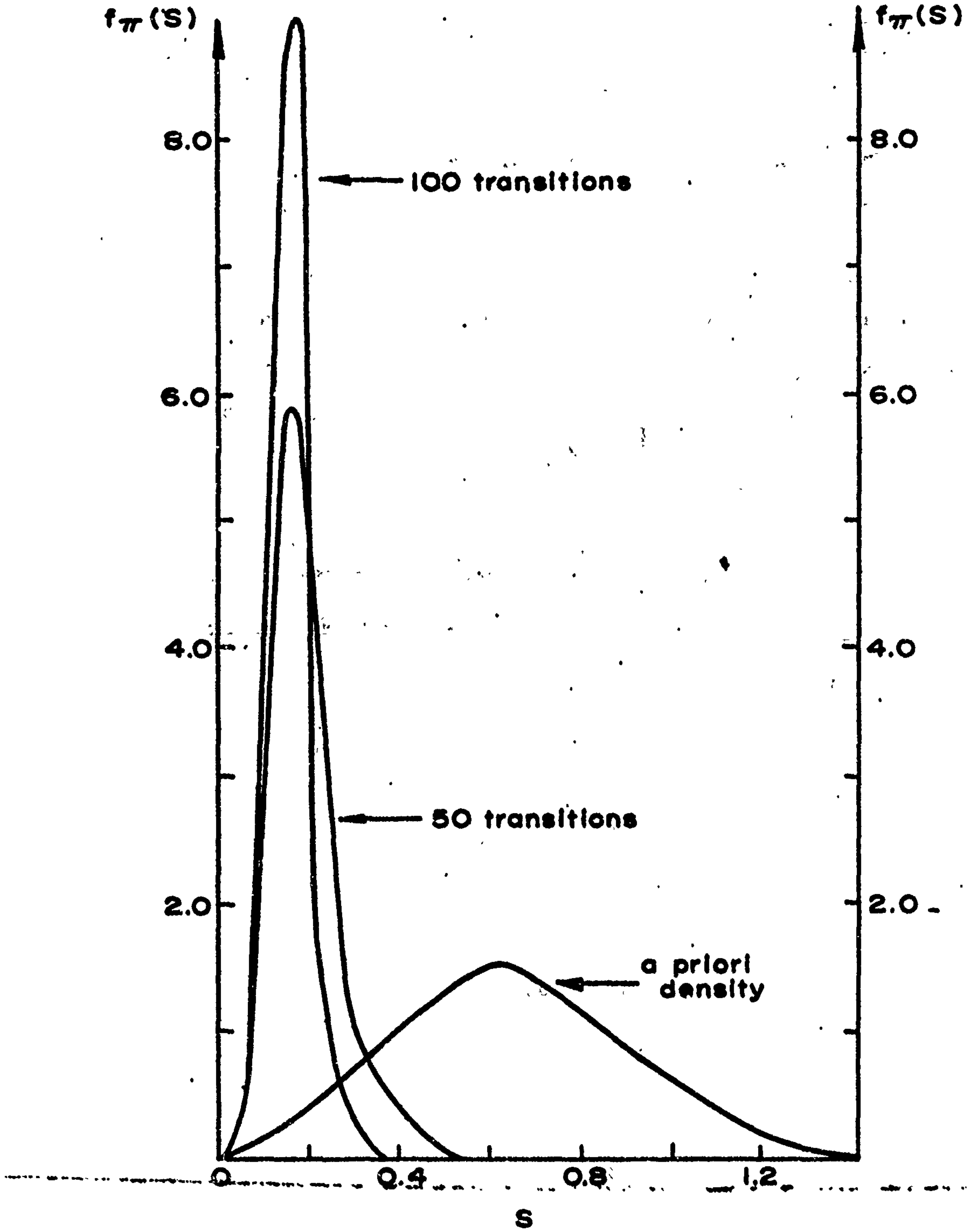


Figure 3.3 Density $f_\pi(s|M+F)$

<u>Example 3.5</u>: Consider another two state Markov chain with matrix P given by

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{bmatrix}$$

Vector $\underline{\pi}$ is

$$\underline{\pi} = (0.9 \quad 0.1)$$

and its representation in T_2 is given by the scalar s where

$$s = 0.1414$$

The a priori density is specified as matrix beta with parameter M,

$$M = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

Matrix F at 50 transitions was

$$F = \begin{bmatrix} 46 & 5 \\ 5 & 2 \end{bmatrix}$$

and at 100 transitions it was

$$F = \begin{bmatrix} 88 & 9 \\ 9 & 2 \end{bmatrix}$$

Density $f_{\pi}(s|M+F)$ is drawn in Figure 3.4. The probability mass is concentrating at $s = 0.1414$.

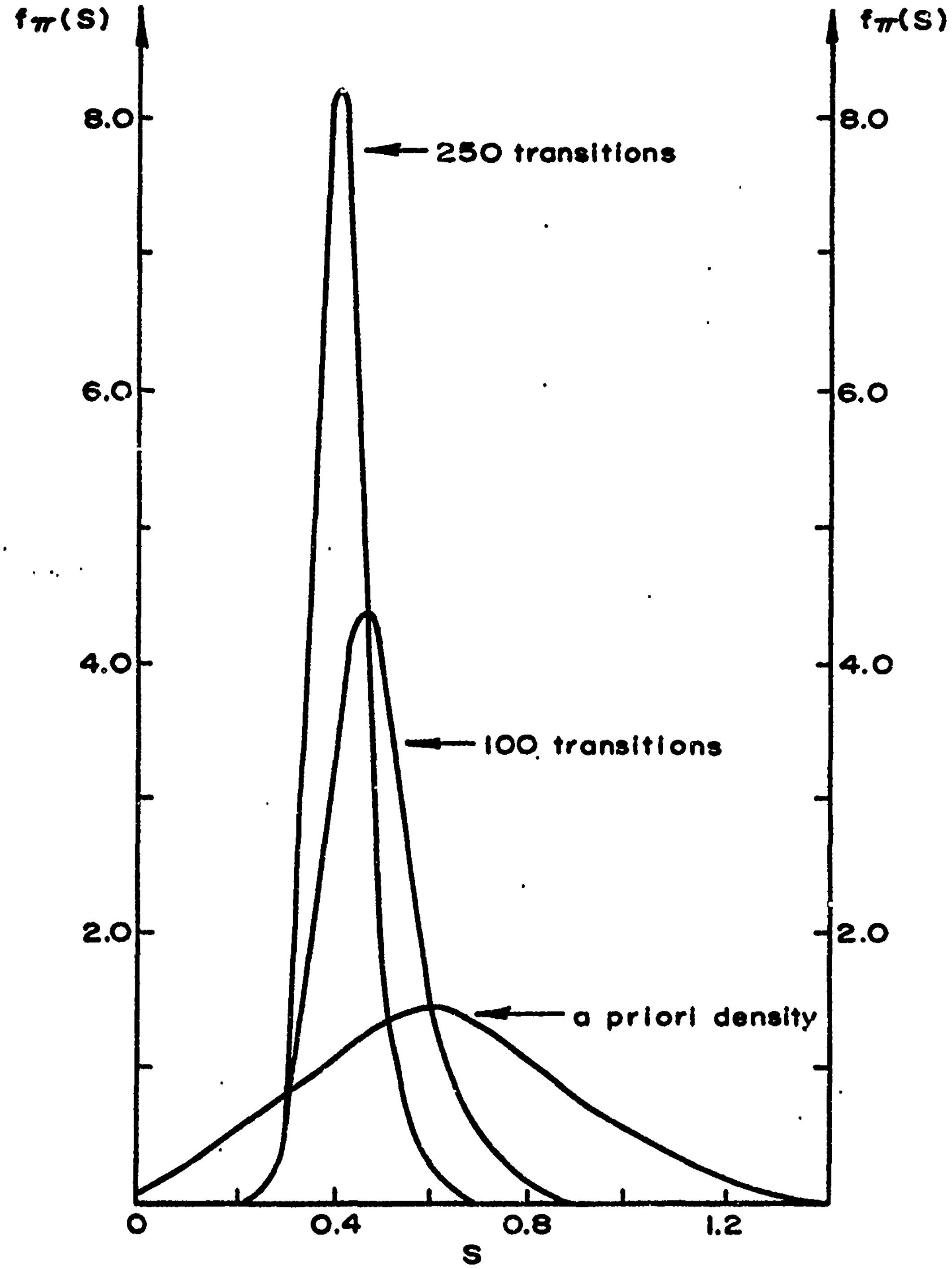


Figure 3.4 Density $f_{\pi}(s|M+F)$

3.6 The Expected Value of the Gain

The expected value of the gain conditioned on the observation $\underline{x}_n$ follows directly from Equation 2.3,

$$E(\Delta \mid \underline{x}_n) = \int_{T_N^N} \langle \underline{w}(\underline{t})U, \underline{r} \rangle f_P(\underline{t} \mid M+F)\, d\underline{t} + \langle \underline{\pi}_o, \underline{r} \rangle \qquad (3.8)$$

Chapter 4

APPLYING BAYESIAN DECISION THEORY

In this chapter there are K decisions. Each decision specifies a unique transition matrix and reward vector. The transition probabilities are unknown. In order to apply Bayesian decision theory the states of nature, a priori density, observation, and loss function have to be specified. With these elements the risk associated with each decision can be computed. The decision maker chooses the decision that minimizes his risk.

4.1 Elements of Bayesian Decision Theory

The states of nature was defined as the set

$$\Omega = \left\{\Lambda^N, \Lambda_2^N, \ldots, \Lambda_K^N\right\}$$

where $\Lambda_i^N = \Lambda^N$ the set of all possible transition matrices under decision i. An element ω of set Ω is

$$\omega = (P_1, P_2, \ldots, P_K)$$

where P_i is the NxN transition matrix under decision i.

The a priori density is defined as the product of matrix beta densities with parameter M^i $i = 1, \ldots, K$.

 Preceding page blank

$$\xi_o(\omega) = f^N_{M\beta}(P_1|M^1)\, f^N_{M\beta}(P_2|M^2) \ldots f^N_{M\beta}(P_K|M^K), \omega \in \Omega$$

The observation $\underline{x}_n$ is given by the sequence,

$$\underline{x}_n = \left\{ \underline{x}^1_{n_1}, \underline{x}^2_{n_2}, \ldots, \underline{x}^K_{n_K} \right\}$$

where

$$n = \sum_{i=1}^{K} n_i$$

and $\underline{x}^i_{n_i}$ is the sequence of states observed under decision i,

$$\underline{x}^i_{n_i} = \left\{ x^i_o, x^i_1, \ldots, x^i_{n_i} \right\}$$

where

$$1 \leq x^i_j \leq K \qquad i = 1, \ldots, K$$

$$j = 1, \ldots, n_i$$

The posteriori density is derived from Equation 1.3,

$$\xi(\omega|\underline{x}_n) = f^N_{M\beta}(P_1|M^1+F^1) \ldots f^N_{M\beta}(P_K|M^K+F^K)$$

The loss function $L(i|\omega)$ was defined in Equation 1.1,

$$L(i|\omega) = \max_j \left\{ < \underline{\sigma}(P_j), \underline{r}^j > - < \underline{\sigma}(P_i), \underline{r}^i > \right\}$$

To select a decision after a finite number of transitions a risk is computed for each decision. The decision maker selects the decision that minimizes the risk.

There are two problems in using this technique to select a decision. First, a sampling strategy has to be specified. The sampling strategy involves the number of transitions recorded under each decision and manner in which one decision is switched to another. Second, a stopping rule has to be specified. State transitions cannot be observed forever. Some rule that indicates when enough information has been collected is needed.

In the case where the decision process makes a finite number of transitions the sampling strategy is crucial. Two goals must be kept in mind when sampling. First, the information gained through sampling should be maximized and second, the payoff should be maximized. In the case where the decision process makes an infinite number of transitions the sampling strategy is designed to maximize the information while neglecting the payoff during the finite sampling period.

In this dissertation the decision process makes an infinite number of transitions. Since the central issue is computing the risk, the sampling strategy adopted here is simply to sample equally under each decision before selecting the risk minimizing decision. A stopping rule is not specified.

The risk is defined as the expected loss,

$$\rho(i) = \int_{\Omega} \max_{j} \left\{ < \underline{\sigma}(P_j), \underline{r}^j > - < \underline{\sigma}(P_i), \underline{r}^i > \right\} \xi(\omega \mid \underline{x}_n)$$

In Chapter 1 it was shown that the risk minimizing decision k* minimizes the function $\eta^i(k)$,

$$\eta^i(k^*) = \min \left\{ \eta^i(k) \right\}$$

where

$$\eta^i(k) = E(\Delta_k \mid \underline{x}_n) - E(\Delta_i \mid \underline{x}_n) \tag{1.6}$$

Substituting Equation 3.8 with decisions into Equation 1.6 gives the desired minimizing function,

$$\eta^i(k) = < \underline{\pi}_o, \underline{r}^k - \underline{r}^i > + \int_{T_N^N} < \underline{w}(\underline{t}) \, U, \underline{r}^k > f_{\underline{p}}(\underline{t} \mid M^k + F^k) \, d\underline{t}$$

$$- \int_{T_N^N} < \underline{w}(\underline{t}) \, U, \underline{r}^i > f_{\underline{p}}(\underline{t} \mid M^i + F^i) \, d\underline{t}, \tag{4.1}$$

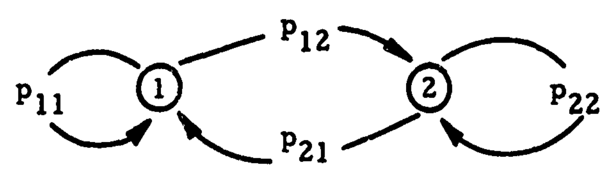
$$k = 1, \ldots, K$$

$$1 \leq i \leq K$$

4.2 Howard's Toymaker Example

The following example was used by Ron Howard [7] to illustrate the procedure used in selecting the decision that maximizes the gain. An example of a Markov decision process can be thought of as the toymaker's process. The toymaker is involved in the novelty toy business. He may be in either of two states. He is in the first state if the toy he is currently producing has found great favor with the

public. He is in the second state if his toy is out of favor. Suppose that when he is in state 1 there is p_{11} percent chance of his remaining in state 1 at the end of the week and, a 1 - p_{11} (=p_{12}) percent chance of a transition to state 2. When he is in state 2 he experiments with new toys, and he may return to state 1 after a week with probability p_{21} or remain unprofitable in state 2 with probability p_{22}. A transition diagram of the system showing the states and transition probabilities in graphical form is



The transition matrix P is given by

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

When the toymaker has a successful toy he earns r_1 units for that week, and if his toy is unsuccessful he earns r_2 units for that week.

Suppose now that the toymaker has other courses of action open to him that will change the probabilities and rewards governing the process. When the toymaker has a successful toy he may use advertising to decrease the chance that the toy will fall from favor. However, because of the advertising cost, the profits to be expected per week will generally be lower. To be specific, suppose that the probability distribution for transitions from state 1 will be

$\underline{p}_1 = (0.8 \;\; 0.2)$ when advertising is employed, and that the corresponding reward will be $r_1 = 4$. The toymaker has two alternatives when he is in state 1: He may use no advertising or he may advertise. These alternatives will be labeled 1 and 2, respectively. Each alternative has its associated reward in state 1 and probability distributions for transitions out of state 1. Alternatives will be indicated by superscript. Thus, for alternative 1 in state 1, $\underline{p}_1^1 = (0.5 \;\; 0.5)$, $r_1^1 = 6$; and for alternative 2 in state 1, $\underline{p}_1^2 = (0.8 \;\; 0.2)$, $r_1^2 = 4$.

There may also be alternatives in state 2 of the system. Increased research expenditures may increase the probability of obtaining a successful toy, but they will also increase the cost of being in state 2. Under alternative 1, a limited research alternative, the probability distribution is $\underline{p}_2^1 = (0.4 \;\; 0.6)$ and the reward is $r_2^1 = -3$. Under the research alternative, alternative 2, the probability and reward distribution is $\underline{p}_2^2 = (0.7 \;\; 0.3)$ and $r_2^2 = -5$.

The alternatives for the toymaker are presented in Table 4.1. A decision is defined as a vector of alternatives in each state. Therefore, there are four decisions. Decision one is defined as alternative 1 in state 1 and alternative 2 in state 2, and so on. Each decision specifies a transition matrix. For example, decision 3 given by alternative 2 in state 1 and alternative 1 in state 2, gives the following transition matrix,

$$P_3 = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}$$

The corresponding steady state probability vector is

$$\underline{\pi}^3 = (0.667 \quad 0.333)$$

and, the gain or expected payoff is

$$\Delta_3 = 2/3(4) + 1/3(-3)$$
$$= 1.67$$

The steady state probability vector and gain for each decision is listed in Table 4.2. If the decision maker has perfect knowledge of the transition probabilities he will select decision 4, the decision that maximizes his expected payoff.

Now, assume that the transition probabilities are unknown. The states of nature are given by

$$\Omega = \left\{\Lambda^2, \ \Lambda^2, \ \Lambda^2, \ \Lambda^2\right\}$$

The a priori density is

$$\xi_o(\omega) = f^2_{M\beta}(P_1 \,|\, M) \ldots f^2_{M\beta}(P_4 \,|\, M)$$

where the "parameter" M is selected as

$$M = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

The Markov chain defined above was simulated on the computer and was observed using the following sampling strategy. Ten transitions

are recorded under each decision. The decisions are switched from decision 1 to decision 2 to decision 3 to decision 4 to decision 1, and so on. The transformed risk function $\pi^1(i)$ is computed from Equation 4.1. The values of $\pi^1(k)$ $k = 1, \ldots, 4$ are plotted in Figure 4.1 every ten transition sequence. The results show that the risk minimizing decision k* is decision 4, the decision that maximizes the gain.

Table 4.1

Transition Vectors and Rewards

State i	Alternative k	Transition Vector p_{i1}^{k}	p_{i2}^{k}	Reward r_i^k
1	1	0.5	0.5	6
1	2	0.8	0.2	4
2	1	0.4	0.6	-3
2	2	0.7	0.3	-5

Table 4.2

Steady State Probabilities and Gains

Decision k	Alternative in State 1	Alternative in State 2	π_1^k	π_2^k	Δ_k
1	1	1	4/9	5/9	1.00
2	1	2	7/12	5/12	1.42
3	2	1	2/3	1/3	1.67
4	2	2	7/9	2/9	2.00

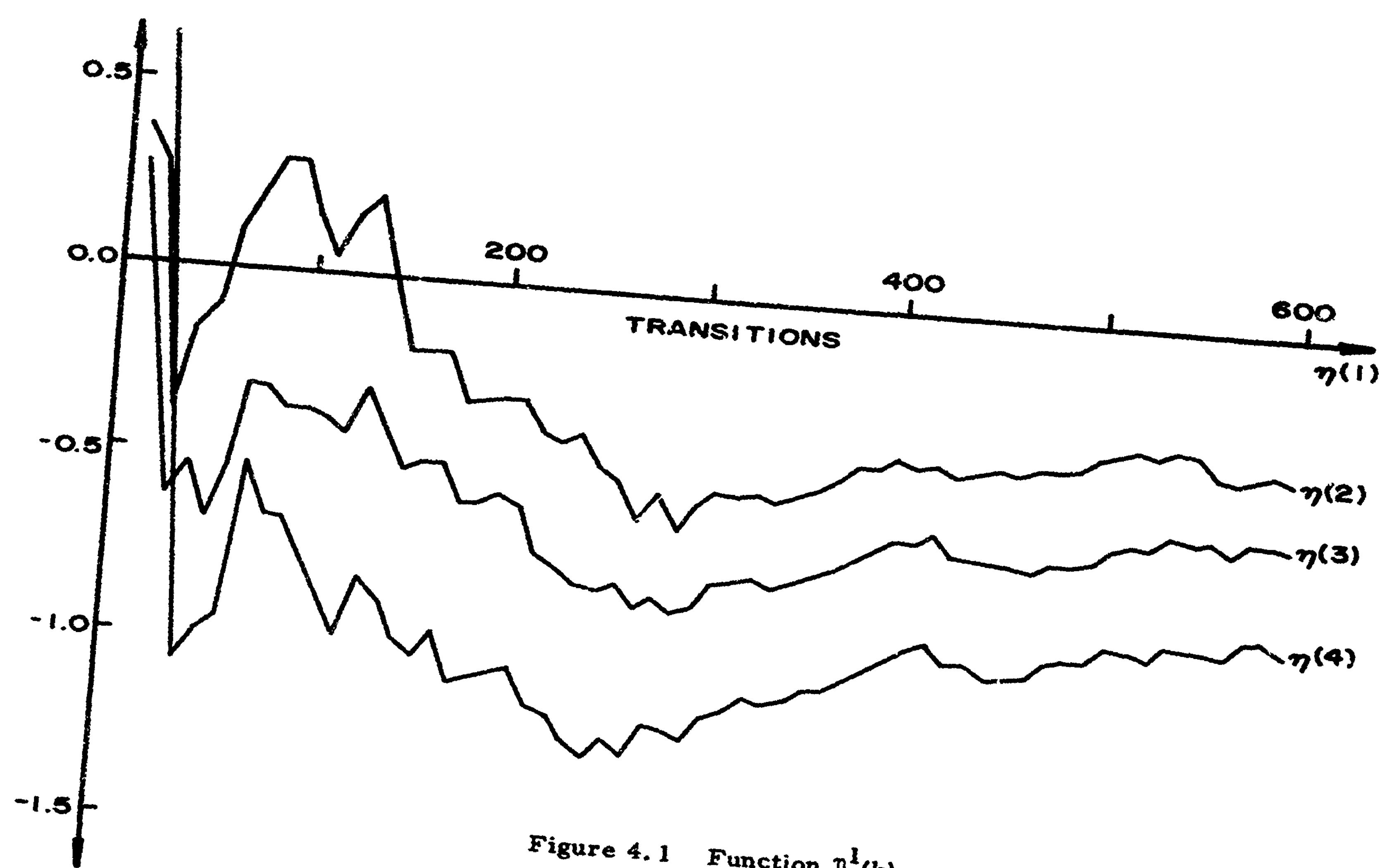


Figure 4.1 Function $\eta^{1}(k)$

Chapter 5

BAYESIAN DECISION THEORY WITH THE UNCERTAINTY OVER THE STEADY STATE PROBABILITIES

Once density $f_P(\underline{t}|M+F)$ is specified, the transformed risk function $\Pi^i(k)$ is evaluated by integrating in the set T_N^N. As N gets large this task becomes increasingly more difficult and time consuming. As an alternative, the a priori density could be placed over the steady state probabilities rather than the transition probabilities. This approach greatly reduces the computational exercise. In this case the states of nature are defined as

$$\Omega = \left\{\Pi_N, \ldots, \Pi_N\right\}$$

and the a priori density is

$$\xi_o(\omega) = f_{V\beta}^N(\underline{\pi}^1|\underline{m}^1)\ldots f_{V\beta}^N(\underline{\pi}^K|\underline{m}^K)$$

where $f_{V\beta}^N(\underline{\pi}^i|\underline{m}^i)$ is the vector beta density with "parameter" $\underline{m}^i$ and is given by $h_i(\underline{p}_i|\underline{m}^i)$ in Equation 3.1 with $\underline{p}_i$ replaced by $\underline{\pi}^i$.

The posteriori density is also vector beta with "parameter" $\underline{m}^i+\underline{f}^i \quad i = 1, \ldots, K$

$$\xi(\omega|\underline{x}_n) = f_{V\beta}^N(\underline{\pi}^1|\underline{m}^1+\underline{f}^1)\ldots f_{V\beta}^N(\underline{\pi}^K|\underline{m}^K+\underline{f}^K)$$

Observation $\underline{x}_n$ is the same as defined in Chapter 4. Vector $\underline{f}^i$ is called the frequency count vector. The j^{th} component of $\underline{f}^i$ is the number of times state j is observed under decision i in sequence $\underline{x}_n$.

The loss function for this case is also the same as defined in Chapter 4. The decision maker chooses the decision k* that minimizes, $\eta^i(k)$. The expected value of the gain $E(\Delta_k | \underline{x}_n)$ is computed as

$$E(\Delta_k | \underline{x}_n) = < E(\underline{\pi}^k | \underline{x}_n), \underline{r}^k >$$

The expected value of $\underline{\pi}^k$ is given by the properties of the matrix beta density in Equation 3.2,

$$E(\underline{\pi}^k | \underline{x}_n) = \frac{1}{\alpha_k} (\underline{m}^k + \underline{f}^k)$$

where

$$\alpha_k = \sum_{j=1}^{N} (m_j^k + f_j^k)$$

Therefore, $\eta^i(k)$ is given by

$$\eta^i(k) = \frac{1}{\alpha_k} < \underline{m}^k + \underline{i}^k, \underline{r}^k > - \frac{1}{\alpha_i} < \underline{m}^i + \underline{f}^i, \underline{r}^i > \qquad (5.1)$$

This expression is easy to evalute. No integration operation is required. Vectors $\underline{m}^i$ $i = 1, \ldots, K$ are specified when the a priori density is defined, and vectors $\underline{f}^i$ $i = 1, \ldots, K$ are defined by the sequence $\underline{x}_n$ observed from the Markov chains.

The critical assumption made in this approach is that the Markov chain being observed is in steady state. However, the Markov chain under observation may not be in steady state. Since the transition probabilities are unknown, the probability distribution over the states is unknown.

Suppose that the Markov chain under consideration has been operating under decision 1 for a large number of transitions. Assume that the probability distribution over the states is $\underline{\pi}^1$, the steady state probability distribution. Now, decision 1 is switched to decision m. The probability distribution over the states after n transition is

$$\underline{\pi}^m(n) = \underline{\pi}^1 P_m^n$$

If observations are recorded before $\underline{\pi}^m(n) \rightarrow \underline{\pi}^m$ then the posteriori distribution over the states of nature will not accurately reflect the knowledge of the steady state probabilities. The question asked is: How many transitions are required before $\underline{\pi}^m(n)$ is "close enough" to $\underline{\pi}^m$ so that observations can be recorded? This question will be answered for special classes of Markov chains in Chapter 6, but under conditions of perfect knowledge of the transition probabilities. Since the transition probabilities are unknown, these results cannot be used. Aside from these theoretical problems, this approach has great practical appeal, especially for Markov chains with large numbers of states.

Example 5.1: The following two state, two decision example is presented to indicate that the approach presented in this Chapter does give good results. The a priori "parameters" $\underline{m}^1$ and $\underline{m}^2$ are given by

$$\underline{m}^1 = \underline{m}^2 = (2 \quad 2)$$

The transition matrices simulated were

$$P_1 = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix}.$$

and the rewards were

$$\underline{r}^1 = (6 \quad -3)$$

$$\underline{r}^2 = (6 \quad -5)$$

The sampling strategy chosen was to sample twenty transitions under each decision. When one decision was switched to another, the first ten transitions were not recorded to allow the probability distribution to near steady state. The transformed risk $\eta^1(i)$ is given by

$$\eta^1(1) = 0$$

$$\eta^1(2) = \frac{1}{\alpha_1} < \underline{m}^1 + \underline{f}^1, \underline{r}^1 > - \frac{1}{\alpha_2} < \underline{m}^2 + \underline{f}^2, \underline{r}^2 >$$

Variable $\eta^1(2)$ is plotted in Figure 5.1. As the number of observations increased, $\eta^1(2)$ approached its true value of -0.42.

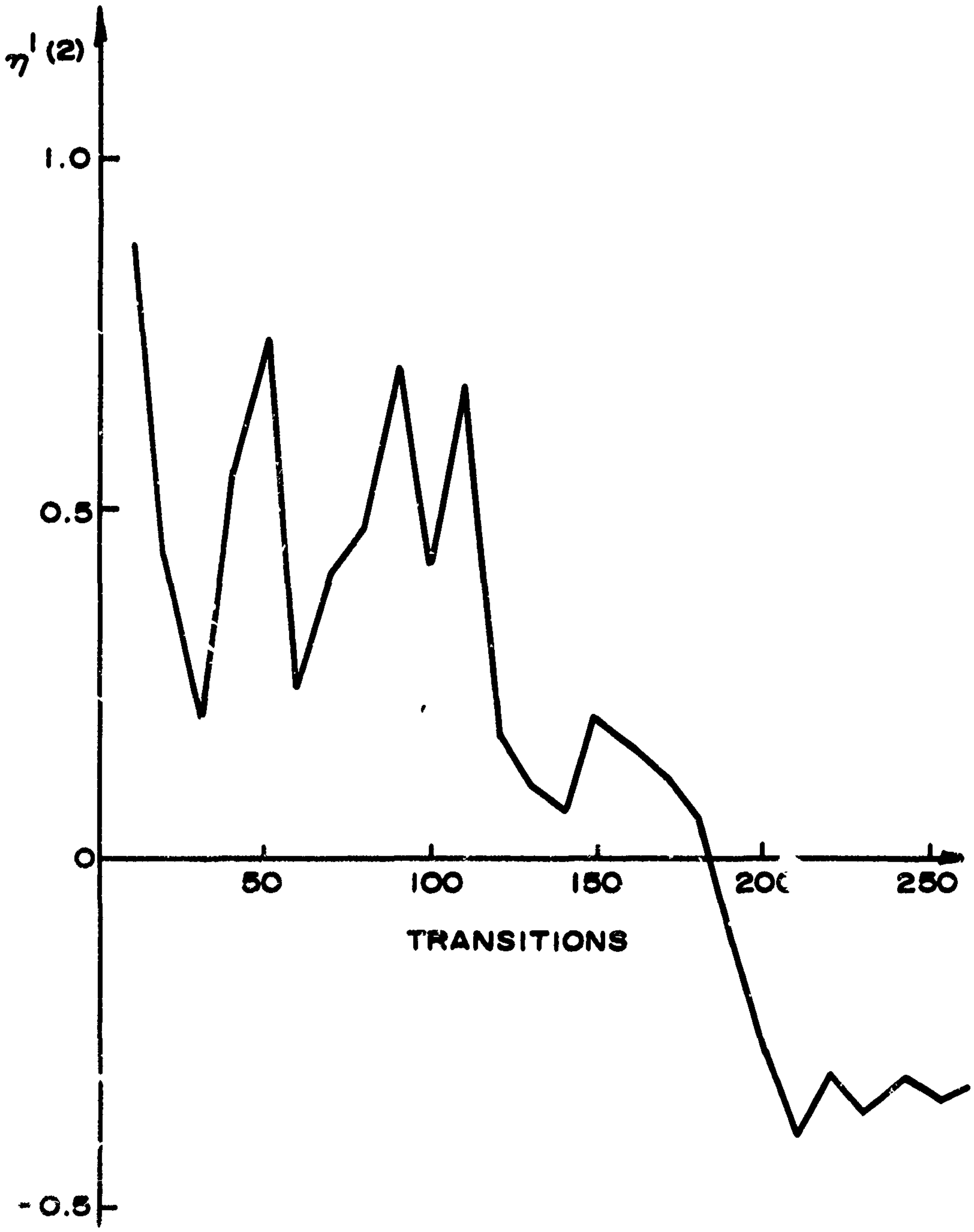


Figure 5.1 Variable $\eta^1(k)$

Chapter 6

CONVERGENCE PROPERTIES OF TWO STATE MARKOV CHAINS

Given an ergodic transition matrix P and an initial state probability vector $\underline{\pi}(o)$ the state probability vector after n transitions, $\underline{\pi}(n)$, is given by

$$\underline{\pi}(n) = \underline{\pi}(o)\, P^n \tag{6.1}$$

As n approaches infinity, vector $\underline{\pi}(n)$ asymptotically approaches the steady state probability vector $\underline{\pi}$. Using sets T_N and T_N^N, the convergence rate of $\underline{\pi}(n)$ to $\underline{\pi}$ for two state Markov chains can be stated explicitly.

In order to determine convergence of $\underline{\pi}(n)$ to $\underline{\pi}$ both probability vectors must be transformed to vectors in T_N. Vector $\underline{\pi}(n)$ can be written as $\underline{\pi}_o + \underline{s}(n)\, U$. For convenience, vector $\underline{\pi}_o$ is defined as the steady state probability vector π. Using this definition

$$\underline{\pi}(n) = \pi + \underline{s}(n)\, U$$

and vector $\underline{\pi}(n) - \pi$ becomes

$$\underline{\pi}(n) - \pi = \underline{s}(n)\, U \tag{6.2}$$

The next step is to evaluate vector $\underline{s}(n)$. From Equation 6.1

$$\begin{aligned}\underline{\pi}(n) &= \underline{\pi} + \underline{s}(n)\, U \\ &= (\pi + \underline{s}(o)\, U)\, P^n\end{aligned}$$

 Preceding page blank

or

$$\underline{s}(n) = \underline{s}(o)\ UP^nV \tag{6.3}$$

Matrix P can be written as

$$P = P_o + SU$$

Similarly, matrix P^n can be written as

$$P^n = P_o + S(n)\ U \tag{6.4}$$

Substituting Equation 6.4 into Equation 6.3 gives

$$\begin{aligned} \underline{s}(n) &= \underline{s}(o)\ UP_oV + \underline{s}(o)\ US(n)\ UV \\ &= \underline{s}(o)\ US(n) \end{aligned} \tag{6.5}$$

Matrix S(n) can be evaluated by manipulating the identity

$$P^{n+1} = P^n\ P \tag{6.6}$$

The i^{th} row of Expression 6.6 satisfies

$$\underline{p}_i^{n+1} = \underline{p}_i^n\ P$$

or

$$\underline{\pi} + \underline{s}_i(n+1)\ U = (\underline{\pi} + \underline{s}_i(n)\ U)P$$

and

$$\underline{s}_i(n+1) = \underline{s}_i(n)\ UPV$$

Therefore

$$S(n+1) = S(n)\ UPV \tag{6.7}$$

Substituting the identity $PV = S + P_o V$ into Equation 6.7 gives

$$
\begin{aligned}
S(n+1) &= S(n)\, US + S(n)\, UP_o V \\
&= S(n)\, US
\end{aligned}
$$

or

$$
\begin{aligned}
S(n+1) &= S(1)\, (US)^n \\
&= S(US)^n
\end{aligned} \tag{6.8}
$$

Substituting Equation 6.8 into Equation 6.5,

$$
\begin{aligned}
\underline{s}(n) &= \underline{s}(o)\, US\, (US)^{n-1} \\
&= \underline{s}(o)\, (US)^n
\end{aligned} \tag{6.9}
$$

Substituting Equation 6.9 into Equation 6.2 gives the desired expression,

$$
\underline{\pi}(n) - \underline{\pi} = \underline{s}(o)\, (US)^n\, U \tag{6.10}
$$

Consider a two state Markov chain. "Matrices" S and U are given by

$$
S = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}
$$

$$
U = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}
$$

Substituting these expressions into Equation 6.10 gives

$$
\begin{aligned}
\underline{\pi}(n) - \underline{\pi} &= \underline{s}(o) < \underline{s}, \underline{u} >^n \underline{u} \\
&= (< \underline{s}, \underline{u} >^n s(o))\, \underline{u}
\end{aligned}
$$

Expression $<\underline{s}, \underline{u}>^n s(o)$ is given by

$$<\underline{s}, \underline{u}>^n s(o) = \left(\frac{s_2 - s_1}{\sqrt{2}}\right)^n <\underline{\pi}(o)-\underline{\pi}, \underline{v}>$$

where

$$v = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad .$$

Therefore, Equation 6.10 for two state Markov chains becomes

$$\underline{\pi}(n) - \underline{\pi} = \left(\frac{s_2 - s_1}{\sqrt{2}}\right)^n <\underline{\pi}(o)-\underline{\pi}, \underline{v}> \underline{u} \tag{6.11}$$

The rate of convergence of $\underline{\pi}(n)$ to π can be determined by finding the integer N such that $||\underline{\pi}(n)-\underline{\pi}||$ is less than some small number ϵ for all $n \geq N$. The norm for two state Markov chains is

$$||\underline{\pi}(n)-\underline{\pi}|| = \left|\frac{s_2 - s_1}{\sqrt{2}}\right|^n |<\underline{\pi}(o)-\underline{\pi}, \underline{v}>|$$

since $||\underline{u}|| = 1$. Letting $||\underline{\pi}(n) - \underline{\pi}|| = \epsilon$ and solving for n results in the expression

$$n = \frac{\log\left[\frac{\epsilon}{|<\underline{\pi}(o)-\underline{\pi}, v>|}\right]}{\log\left[\left|\frac{s_2 - s_1}{\sqrt{2}}\right|\right]} \tag{6.12}$$

The expression for $\underline{\pi}(n)$ can be used to investigate the convergence of transition matrix P^n to the limiting matrix P^∞. For two state Markov chains, the two transition vectors of P^n, $\underline{p}_1^n$, $\underline{p}_2^n$, are

written as

$$\underline{p}_1^n = \underline{\pi} + s_1(n)\, U$$

$$\underline{p}_2^n = \underline{\pi} + s_2(n)\, U$$

Convergence of P^n to P^∞ in set Λ^2 is equivalent to the convergence of vector $\underline{t}(n)$ to $\underline{\sigma} = (S, S)$ in set T_2^2

$$\underline{t}(n) = (s_1(n),\ s_2(n))$$

$$s = \langle (\underline{\pi} - \underline{\pi}_o), \underline{v} \rangle = 0$$

where the expressions for scalars $s_1(n)$ and $s_2(n)$ are derived rather simply from Equation 6.9,

$$s_1(n) = s_1(o)\ (US)^n \qquad (6.13)$$

$$s_2(n) = s_2(o)\ (US)^n \qquad (6.14)$$

where

$$s_1(o) = s_1 = \langle (\underline{p}_1 - \underline{\pi}), \underline{v} \rangle \qquad (6.15)$$

$$s_2(o) = s_2 = \langle (\underline{p}_2 - \underline{\pi}), \underline{v} \rangle \qquad (6.16)$$

Equations 6.13, 6.14, 6.15, and 6.16 combine to form vector $\underline{t}(n)$,

$$\underline{t}(n) = \underline{t}(US)^n$$

or

$$\underline{t}(n) = \left(\frac{s_2 - s_1}{\sqrt{2}} \right)^n \underline{t} \qquad (6.17)$$

Example 6.1: Consider a two state Markov chain with transition matrix

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

Vector $\underline{\pi}_o$ is defined by

$$\underline{\pi}_o = \underline{\pi} = (1/3,\ 2/3)$$

Therefore, set T_2 is defined by

$$T_2 = \left\{ s \in E' \mid -\frac{2\sqrt{2}}{3} \leq s \leq \sqrt{2}/3 \right\}$$

Vector $\underline{t}(n)$ in Equation 6.17 is

$$\underline{t}(n) = \left(\frac{1}{4}\right)^n \left(\frac{\sqrt{2}}{2},\ \frac{3\sqrt{2}}{4}\right)$$

Chapter 7

SUMMARY AND RECOMMENDATIONS

The development presented in this dissertation is summarized in this Chapter and topics for future research are proposed. The objective of this research was to apply Bayesian decision theory to the Markov decision problem with unknown transition probabilities. The Markov chain under consideration had N states and made an infinite number of transitions. There were K decisions. Each decision i specified a transition matrix P_i and reward vector $\underline{r}^i$. The states of nature Ω were defined by

$$\Omega = \left\{ \Lambda_1^N, \ldots, \Lambda_K^N \right\}$$

The a priori density over Ω, $\xi_o(\omega)$ was matrix beta. Then the posteriori density $\xi(\omega | \underline{x}_n)$ was computed from the a priori density, observation $\underline{x}_n$, and Bayes formula. The loss function $L(i | \omega)$ specified the loss per transition when the state of nature was ω and decision i was selected. The risk function $\rho(i)$ was defined as the expected loss

$$\rho(i) = \int_\Omega L(i | \omega) \, \xi(\omega | \underline{x}_n) \, d\omega$$

The risk minimizing decision k* was shown to also minimize function $\eta^i(k)$ where

$$\eta^i(k) = E(\Delta_k | \underline{x}_n) - E(\Delta_i | \underline{x}_n)$$

Therefore, decision k* satisfies

$$\eta^i(k^*) = \min_k \left\{ E(\Delta_k | \underline{x}_n) - E(\Delta_i | \underline{x}_n) \right\}$$

where

$$E(\Delta_k | \underline{x}_n) = \int_{T_N^N} \langle \underline{w}(\underline{t})U, \underline{r}^k \rangle f_p(\underline{t} | M^k + F^k)\, dt + \langle \underline{\pi}_o, \underline{r}^k \rangle$$

$$T_N^N = T_N \times \ldots \times T_N \quad \text{(N times)}$$

$$T_N = \left\{ \underline{s} \in E^{N-1} | \underline{s}U = \underline{p} - \underline{\pi}_o,\ \underline{p} \in \Pi_N \right\}$$

$$\underline{\pi}_o = \underline{e}_i$$

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} \underline{e}_2 \\ \vdots \\ \underline{e}_N \end{bmatrix} - P_o$$

$$V = \left[\underline{e}_2\ \underline{e}_3 \cdots \underline{e}_N \right]$$

$$P_o = \begin{bmatrix} \underline{\pi}_o \\ \vdots \\ \underline{\pi}_o \end{bmatrix}$$

$$\underline{w}(\underline{t}) = \underline{\pi}_o S(I - US)^{-1}$$

$$\underline{t} = (\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_N)$$

$$S = \begin{bmatrix} \underline{s}_1 \\ \vdots \\ \underline{s}_N \end{bmatrix}$$

$$f_p(\underline{t} | M^k + F^k) = \prod_{i=1}^{N} \frac{1}{\det [J]} h_i(\underline{\pi}_o + \underline{s}_i U | \underline{m}_i^k + \underline{f}_i^k)$$

$$\underline{m}_i^k = i^{th} \text{ row of } M^k$$

$$\underline{f}_i^k = i^{th} \text{ row of } F^k$$

$$h_i(\underline{\pi}_o + \underline{s}_i U \mid \underline{m}_i^k + \underline{f}_i^k) = k(\underline{m}_i^k + \underline{f}_i^k) \prod_{j=1}^{N} (\underline{\pi}_{oj} + \langle \underline{s}_i, \underline{U}_j \rangle)^{m_{ij}^k + f_{ij}^k - 1}$$

The objectives of this research were met. A simple solution to the Markov decision problem with uncertainty has been derived. However, the problem is not completely solved. The following is a list of topics that require future research.

1. The problem of selecting an "optimal" sampling strategy and an "optimal" stopping rule is a candidate for future research. Martin discusses this problem at length. However, his results do not appear to be amenable to a practical application. It is possible that this problem could be successfully analyzed by using the framework developed in this dissertation.

2. In this paper the Markov decision process under consideration was an infinite stage process. The case where the Markov chain makes a finite number of transitions should be analyzed. Here the sampling strategy will be of primary importance because two goals will be present during the life of the process. One is to maximize the payoff and the other is to maximize the information. In

the infinite stage process the payoff was not an issue since the Markov chain would make an infinite number of transitions after a decision was selected.

3. The case where the uncertainty is placed over the steady state probabilities was discussed in Chapter 5. This approach simplifies the computations considerably. However, errors are present because the decision maker does not know whether the Markov chain is in steady state. If some means of approximating the time when steady state is "nearly" reached were found, this approach might be more practical for a Markov chain with a large number of states.

BIBLIOGRAPHY

1. Bellman, Richard, Introduction to matrix analysis, New York, McGraw-Hill, 1960.

2. Bhat, B.R., "Bayes solution of sequential decision problem for Markov dependent observations," Ann. Math. Stat., 35 1965 .

3. DeGroot, Morris H., Optimal statistical decisions, New York, McGraw-Hill, 1970.

4. Ferguson, Thomas S., Mathematical statistics, New York, Academic Press, 1967.

5. Halmos, Paul R., Finite-dimensional vector spaces, New Jersey, D. Van Nostrand, 1958.

6. Hoffman, A.J. and Karp, R.M., "On non-terminating stochastic Games," Management Science, Vol. 12, No. 5, January, 1968.

7. Howard, Ronald A., Dynamic programming and Markov processes, Massachusetts, The M.I.T. Press, 1960.

8. Lee, T.C., Judge, G.G., and Cain, R.L., "A sampling study of the properties of estimators of transition probabilities," Management Science, Vol. 2, No. 9, May, 1966.

9. Martin, J.J., Bayesian decision problems and Markov chains, New York, John Wiley & Sons, 1967.

10. Papoulis, Athanasios, Probability, random variables, and stochastic processes, New York, McGraw-Hill, 1965.

11. Pollatschek, M.A. and Avi-Itzhak, B., "Algorithms for stochastic games with geometric interpretation," Management Science, Vol. 15, No. 7, March, 1969.

12. Schweitzer, Paul J., "Perturbation theory and finite Markov chains," J. Appl. Prob., 5, 401-413 1968 .

13. Shapley, L., "Stochastic games," Proc. Nat. Acad. Sci., 39, pp. 1095-1100.

14. Theil, H. and Rey, Guido, "A quadratic programming approach to the estimation of transition probabilities," Management Science, Vol. 2, No. 9, May, 1966.